MARYGROVE COLLEGE
LIBRARY
DETROIT, MICH.

THE ACCOUNTING REVIEW

VOLUME 85, No. 4 JULY 2010

A FORUM ON INDIVIDUAL DIFFERENCES IN ACCOUNTING BEHAVIOR

STEVEN J. KACHELMEIER
Introduction (Editorial)

LINDA SMITH BAMBER, JOHN (XUEFENG) JIANG, and ISABEL YANYAN WANG
What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure

SCOTT D. DYRENG, MICHELLE HANLON, and EDWARD L. MAYDEW
The Effects of Executives on Corporate Tax Avoidance

OTHER ARTICLES

MARK BAGNOLI and SUSAN G. WATTS
Oligopoly, Disclosure, and Earnings Management

ANNE BEATTY, SCOTT LIAO, and JOSEPH WEBER
Financial Reporting Quality, Private Information, Monitoring, and the Lease-versus-Buy Decision

ROBERT M. BOWEN, ANDREW C. CALL, and SHIVA RAJGOPAL
Whistle-Blowing: Target Firm Characteristics and Economic Consequences

JOSEPH F. BRAZEL, TINA D. CARPENTER, and J. GREGORY JENKINS
Auditors' Use of Brainstorming in the Consideration of Fraud: Reports from the Field

YUN FAN, ABHIJIT BARUA, WILLIAM M. CREADY, and WAYNE B. THOMAS
Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items

KARIM JAMAL and HUN-TONG TAN
Joint Effects of Principles-Based versus Rules-Based Standards and Auditor Type in Constraining Financial Managers' Aggressive Reporting

ROBERT J. RESUTEK
Intangible Returns, Accruals, and Return Reversal: A Multiperiod Examination of the Accrual Anomaly

CHANG JOON SONG, WAYNE B. THOMAS, and HAN YI
Value Relevance of FAS No. 157 Fair Value Hierarchy Information and the Impact of Corporate Governance Mechanisms

YAN SUN
Do MD&A Disclosures Help Users Interpret Disproportionate Inventory Increases?

DAN WEISS
Cost Behavior and Analysts' Earnings Forecasts

CURRENT YR/VOL

Marygrove College Library
8425 West McNichols Road
Detroit, MI 48221

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2009–2010 Executive Committee of the Association are as follows:

President	Nancy A. Bagranoff, Old Dominion University
President-Elect	Kevin D. Stocks, Brigham Young University
Past President	Susan F. Haka, Michigan State University
Vice President	Ira Solomon, University of Illinois at Urbana–Champaign
Vice President–Education	Belverd E. Needles, Jr., DePaul University
Vice President–Finance	Christopher J. Wolfe, Texas A&M University
Vice President–Finance Elect	Stacy E. Kovar, Kansas State University
Vice President–International	Kazuo Hiramatsu, Kwansei Gakuin University
Vice President–Professional Relations	Robert H. Colson, Grant Thornton, LLP
Vice President–Publications	Jean C. Bedard, Bentley University
Vice President–Research	Gregory B. Waymire, Emory University
Vice President–Sections and Regions	Bruce K. Behn, The University of Tennessee

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive *Accounting Education News*, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review*, *Accounting Horizons*, and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues include online selection of one, two, or all three of the following Association-wide journals: *The Accounting Review*, *Accounting Horizons*, and *Issues in Accounting Education*. (This does not include Section journals.)

Select one Association-wide journal	U.S. $180.00	with Print Option $230.00
Select two Association-wide journals	U.S. $190.00	with Print Option $240.00
All three Association-wide journals	U.S. $200.00	with Print Option $250.00

Student membership:
The print option is not available for student members.

Select one Association-wide journal	U.S. $25.00
Select two Association-wide journals	U.S. $35.00
All three Association-wide journals	U.S. $45.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Auditing; Diversity; Financial Accounting and Reporting; Gender Issues and Worklife Balance; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Strategic and Emerging Technologies; Teaching, Learning and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to tar@mccombs.utexas.edu. If it is necessary to send any hard copy correspondence, the mailing address is Mary Capps, Editorial Assistant, *The Accounting Review*, The University of Texas at Austin, Department of Accounting, CBA 4M.202, 1 University Station B6400, Austin, TX 78712-0211 • Telephone: (512) 471-5318 • Fax: (512) 471-3904.

The submission fee of $200.00 for members of the AAA or $400.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review is indexed in the Social Science Citation Index which is included in Web of Science.

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rates of $375 for a print copy or at $425 for a print and electronic copy. The member subscription rate for a hard-copy is $30. Electronic access is included with membership dues. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review*, 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2010 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Senior Editor

Steven J. Kachelmeier

Editors

Shannon W. Anderson
J. Harry Evans III
James E. Hunton
Kathryn Kadous
Sanjay G. Kallapur
Ranjani Krishnan
Laureen A. Maines
D. Paul Newman
Thomas C. Omer
Wayne B. Thomas
Mark A. Trombley
Beverly R. Walther
Paul A. Zarowin

VOL. 85 JULY 2010 No. 4

A FORUM ON INDIVIDUAL DIFFERENCES IN ACCOUNTING BEHAVIOR

STEVEN J. KACHELMEIER
Introduction (Editorial)

LINDA SMITH BAMBER, JOHN (XUEFENG) JIANG, and ISABEL YANYAN WANG
What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure 1131

SCOTT D. DYRENG, MICHELLE HANLON, and EDWARD L. MAYDEW
The Effects of Executives on Corporate Tax Avoidance 1163

OTHER ARTICLES

MARK BAGNOLI and SUSAN G. WATTS
Oligopoly, Disclosure, and Earnings Management 1191

ANNE BEATTY, SCOTT LIAO, and JOSEPH WEBER
Financial Reporting Quality, Private Information, Monitoring, and the Lease-versus-Buy Decision 1215

ROBERT M. BOWEN, ANDREW C. CALL, and SHIVA RAJGOPAL
Whistle-Blowing: Target Firm Characteristics and Economic Consequences 1239

JOSEPH F. BRAZEL, TINA D. CARPENTER, and J. GREGORY JENKINS
Auditors' Use of Brainstorming in the Consideration of Fraud: Reports from the Field 1273

YUN FAN, ABHIJIT BARUA, WILLIAM M. CREADY, and WAYNE B. THOMAS
Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items 1303

KARIM JAMAL and HUN-TONG TAN
Joint Effects of Principles-Based versus Rules-Based Standards and Auditor Type in Constraining Financial Managers' Aggressive Reporting 1325

ROBERT J. RESUTEK
Intangible Returns, Accruals, and Return Reversal: A Multiperiod Examination of the Accrual Anomaly 1347

CHANG JOON SONG, WAYNE B. THOMAS, and HAN YI
Value Relevance of FAS No. 157 Fair Value Hierarchy Information and the Impact of Corporate Governance Mechanisms 1375

YAN SUN
Do MD&A Disclosures Help Users Interpret Disproportionate Inventory Increases? **1411**
DAN WEISS
Cost Behavior and Analysts' Earnings Forecasts **1441**

BOOK REVIEWS, Stephen A. Zeff, Editor

Robert Bloomfield and Kristina Rennekamp
Experimental Research in Financial Reporting: From the Laboratory to the Virtual World ..... KATHRYN KADOUS **1473**
Robert S. Kaplan and David P. Norton
The Execution Premium: Linking Strategy to Operations for Competitive Advantage ..... SHAHID ANSARI **1475**
Stephen Zeff (Editor)
Principles before Standards: The ICAEW's 'N Series' of Recommendations on Accounting Principles 1942–1969 ..... CHRISTOPHER NAPIER **1477**
Errata **1479**
Editorial Policy and Style Information **1481**

Editorial Staff of *The Accounting Review*
2009–2010

SENIOR EDITOR

Steven J. Kachelmeier, The University of Texas at Austin

EDITORS

Shannon W. Anderson, University of California, Davis
J. Harry Evans III, University of Pittsburgh
James E. Hunton, Bentley University
Kathryn Kadous, Emory University
Sanjay G. Kallapur, Indian School of Business
Ranjani Krishnan, Michigan State University
Laureen A. Maines, Indiana University
D. Paul Newman, The University of Texas at Austin
Thomas C. Omer, Texas A&M University
Wayne B. Thomas, University of Oklahoma
Mark A. Trombley, The University of Arizona
Beverly R. Walther, Northwestern University
Paul A. Zarowin, New York University

EDITORIAL ASSISTANT

Mary C. Capps, The University of Texas at Austin

RESEARCH ASSISTANT

Tracie Majors, The University of Texas at Austin

American Accounting Association

EXECUTIVE DIRECTOR

Tracey E. Sutherland

COMMUNICATIONS DEPARTMENT

Beverly Harrelson
Diane Hazard
Lisa Habblitz
Lindsay Wharton
Kathy Casper
Judy Cothern
Sonja Langley
Jan Kovarik (subcontracted)
Copyediting Team (subcontracted)

OTHER AAA STAFF

Stephen Shepard
Dierdre Harris
David Boynton
Alex Metz
Pat Stein
Debora Gardner
Peggy Turczyn
Barbara Gutierrez
Barbara Brady
Arlene Oliver
Kathleen Messenger
Roland LaTulip
Diane Leger
Man Fong
Jodie M. Morrill
Michelle Bowles

Editorial Advisory and Review Board

Anwer S. Ahmed, Texas A&M University
Ashiq Ali, The University of Texas at Dallas
Anil Arya, The Ohio State University
Benjamin C. Ayers, University of Georgia
William R. Baber, Georgetown University
Stephen P. Baginski, University of Georgia
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Linda Smith Bamber, University of Georgia
Orie E. Barron, The Pennsylvania State University
Anne L. Beatty, The Ohio State University
Jean C. Bedard, Bentley University
Robert J. Bloomfield, Cornell University
Jennifer L. Blouin, University of Pennsylvania
Mark T. Bradshaw, Boston College
Lawrence D. Brown, Georgia State University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
Joseph V. Carcello, The University of Tennessee
Mary Ellen Carter, Boston College
Paul K. Chaney, Vanderbilt University
Qiang Cheng, University of Wisconsin–Madison
Michael B. Clement, The University of Texas at Austin
William M. Cready, The University of Texas at Dallas
Antonio Davila, IESE Business School–University of Navarra
Patricia M. Dechow, University of California, Berkeley
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Sunil Dutta, University of California, Berkeley
Leslie G. Eldenburg, The University of Arizona
Patricia M. Fairfield, Georgetown University
David B. Farber, University of Missouri
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
James R. Frederickson, Melbourne Business School
Frank B. Gigler, University of Minnesota
Cristi A. Gleason, The University of Iowa
Johnathan C. Glover, Carnegie Mellon University
Steven M. Glover, Brigham Young University
David A. Guenther, University of Oregon
Luzi Hail, University of Pennsylvania
Jeffrey Hales, Georgia Institute of Technology
John R. M. Hand, The University of North Carolina at Chapel Hill
R. Lynn Hannan, Georgia State University
Gilles Hilary, INSEAD
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
Chris E. Hogan, Michigan State University
Ole-Kristian Hope, University of Toronto
Richard W. Houston, The University of Alabama
Raffi J. Indjejikian, University of Michigan
Christopher D. Ittner, University of Pennsylvania
Debra C. Jeter, Vanderbilt University
Karla M. Johnstone, University of Wisconsin–Madison
Bin Ke, The Pennsylvania State University
S. Jane Kennedy, University of Washington
Inder K. Khurana, University of Missouri
Michael D. Kimbrough, Harvard University

Ronald R. King, Washington University in St. Louis
April Klein, New York University
W. Robert Knechel, University of Florida
Lisa L. Koonce, The University of Texas at Austin
Susan D. Krische, University of Illinois at Urbana–Champaign
Jayanthi Krishnan, Temple University
William N. Lanen, University of Michigan
David F. Larcker, Stanford University
Reuven Lehavy, University of Michigan
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Joan L. Luft, Michigan State University
Anne M. Magro, George Mason University
Carol Marquardt, Baruch College–CUNY
Dawn A. Matsumoto, University of Washington
Edward Maydew, The University of North Carolina at Chapel Hill
Brian W. Mayhew, University of Wisconsin–Madison
Sarah E. McVay, University of Utah
Kenneth A. Merchant, University of Southern California
Gregory S. Miller, University of Michigan
Lillian F. Mills, The University of Texas at Austin
Brian Mittendorf, The Ohio State University
Partha S. Mohanram, Columbia University
Steven J. Monahan, INSEAD
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Arkansas
Venky Nagar, University of Michigan
Terry L. Neal, The University of Tennessee
Karen K. Nelson, Rice University
Edward F. O'Donnell, Illinois State University
Per M. Olsson, Duke University
Chul W. Park, The University of Hong Kong
Mark E. Peecher, University of Illinois at Urbana–Champaign
Sean A. Peffer, University of Kentucky
Kathy Ruby Petroni, Michigan State University
John D. Phillips, University of Connecticut
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, Stanford University
Marlene A. Plumlee, University of Utah
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shiva Rajgopal, Emory University
Scott Richardson, BlackRock Inc.
John R. Robinson, The University of Texas at Austin
Darren T. Roulstone, The Ohio State University
Katherine Schipper, Duke University
Karen L. Sedatole, Michigan State University
Catherine M. Shakespeare, University of Michigan
Lakshmanan Shivakumar, London Business School
K. (Shiva) Sivaramakrishnan, University of Houston
Hollis A. Skaife, University of Wisconsin–Madison
Richard G. Sloan, University of California, Berkeley
J. Reed Smith, Indiana University at Indianapolis
Mark T. Soliman, University of Washington
Geoffrey B. Sprinkle, Indiana University
Mary Harris Stanford, Texas Christian University
Phillip C. Stocken, Dartmouth College
K. R. Subramanyam, University of Southern California
Hun-Tong Tan, Nanyang Technological University
Kristy L. Towry, Emory University
Senyo Tse, Texas A&M University
Jennifer Wu Tucker, University of Florida
Irem Tuna, London Business School
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Michael S. Wilkins, Texas A&M University
Richard H. Willis, Vanderbilt University
Arnold M. Wright, Northeastern University
Teri Lombardi Yohn, Indiana University
Xiao-Jun Zhang, University of California, Berkeley
Mark F. Zimbelman, Brigham Young University

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1959–1962	Robert K. Mautz
1962–1965	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William Kinney
1990–1994	A. Rashad Abdel-khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber
2002–2005	Terry Shevlin
2005–2008	Dan S. Dhaliwal

EDITORIAL

Introduction to a Forum on Individual Differences in Accounting Behavior

Steven J. Kachelmeier
The University of Texas at Austin

This issue of *The Accounting Review* features two related archival studies, both accepted independently by different editors, on the influence of individual managers on different behaviors that are of interest to accountants. Bamber et al. (2010) focus on the effects of individual managers on several properties of management earnings forecasts. Dyreng et al. (2010) focus on the effects of individual managers on tax-avoidance strategies. Although they examine different aspects of management behavior, both studies adapt a technique developed by Bertrand and Schoar (2003) that tracks managers across different firms, thereby disentangling manager effects from firm effects.

From my perspective, both as an editor and as a behavioral experimentalist, a prominent contribution of these two studies is that they help to narrow the gap that too often exists between archival accounting research and so-called "behavioral" accounting research that we typically associate with experiments. For example, Bamber et al. (2010) draw, in part, on a review article by Hirst et al. (2008) regarding management earnings forecasts, which in turn integrates both archival and experimental research perspectives. Indeed, I have often wondered why we tend to associate the adjective "behavioral" with experimental research, although I suppose the emergence of "behavioral finance" as a mainstream archival research area has made inroads in reinforcing that both finance and accounting are ultimately involved with human behavior, reflecting human information processing limitations, biases, and idiosyncrasies (also see Shleifer 2000).

More fundamentally, Bamber et al. (2010) and Dyreng et al. (2010) remind us that firms do not make decisions. Rather, people make decisions, and those decisions are shaped by the personalities of those involved. From a research perspective, one is tempted to build on that premise by identifying characteristics such as gender, education, ethnicity, military service, and a host of other measures that might influence manager styles. On this point, however, I found it of considerable interest that, notwithstanding the spirited attempt in both studies to identify such characteristics, idiosyncratic variation appears to dominate the variation that can be nicely categorized by measurable forces that shape manager personalities.

In the final analysis, it would appear that managers, like other humans, are richly varied, such that the noise term in regressions involving "firm" behavior will likely continue to be larger than researchers would like to see. I applaud the authors of both studies in this forum for helping us to peer inside the noise term in what seems to be the beginnings of a fruitful avenue for future accounting research of archival, experimental, and other varieties.

Published Online: June 2010

REFERENCES

Bamber, L. S., J. X. Jiang, and I. Y. Wang. 2010. What's my style? The influence of top managers on voluntary corporate financial disclosure. *The Accounting Review* 85 (July): 1131–1162.

Bertrand, M., and A. Schoar. 2003. Managing with style: The effects of managers on firm policies. *The Quarterly Journal of Economics* 68 (November): 1169–1208.

Dyreng, S. D., M. Hanlon, and E. L. Maydew. 2010. The effects of executives on corporate tax avoidance. *The Accounting Review* 85 (July): 1163–1189.

Hirst, D. E., L. Koonce, and S. Venkataraman. 2008. Management earnings forecasts: A review and framework. *Accounting Horizons* 22 (September): 315–338.

Shleifer, A. 2000. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford, UK: Oxford University Press.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1131–1162

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1131

What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure

Linda Smith Bamber
The University of Georgia
John (Xuefeng) Jiang
Isabel Yanyan Wang
Michigan State University

ABSTRACT: Financial economics has posited a limited role for idiosyncratic noneconomic manager-specific influences, but the strategic management literature suggests such individual influences can affect corporate outcomes. We investigate whether individual managers play an economically significant role in their firms' voluntary financial disclosure choices. Tracking managers across firms over time, we find top executives exert unique and economically significant influence (manager-specific fixed effects) on their firms' voluntary disclosures, incremental to known economic determinants of disclosure, and firm- and time-specific effects. Managers' unique disclosure styles are associated with observable demographic characteristics of their personal backgrounds: managers promoted from finance, accounting, and legal career tracks, managers born before World War II, and those with military experience develop disclosure styles displaying certain conservative characteristics; and managers from finance and accounting and those with military experience favor more precise disclosure styles. These plausible associations confirm that our estimated manager-specific fixed effects capture systematic long-lived differences in managers' unique disclosure styles.

Keywords: *voluntary disclosure; management earnings forecasts; individual differences; upper echelons theory.*

Data Availability: *Data are publicly available from the sources identified in the text.*

We are grateful for constructive comments from participants in the 2008 Midwest Accounting Research Conference and the 2008 Southeast Summer Accounting Research Colloquium, as well as workshop participants at the University of Connecticut, Florida State University, Norwegian School of Economics and Business Administration, and the Turku School of Economics. Special thanks to Ben Ayers, Mark Bagnoli, Michael Bamber, Larry Brown, Daniel Feldman, Pat Hopkins, Steve Kachelmeier, Grace Pownall, K. Ramesh, Hannu Schadewitz, Jack Stecher, Wayne Thomas, Eric Yeung, two anonymous reviewers, and especially Marilyn Johnson for detailed comments and in-depth discussions that helped us improve this paper. Professor Bamber is grateful for financial support from the J.M. Tull School of Accounting and the Terry College of Business. Professors Jiang and Wang are grateful for financial support from the Department of Accounting and Information Systems at Michigan State University.

Editor's note: Accepted by Wayne Thomas.

Submitted: August 2008
Accepted: January 2010
Published Online: June 2010

I. INTRODUCTION

We investigate two related research questions: (1) Do top managers exhibit unique and economically significant individual-specific styles in voluntary corporate financial disclosure? (2) If so, are managers' unique experiences, as evidenced by observable demographic characteristics of their personal backgrounds, associated with cross-sectional differences in their overall disclosure styles?

Published theoretical and archival research in economics, finance, and accounting has largely posited a limited role for idiosyncratic noneconomic manager-specific influences in explaining cross-sectional differences in corporate choices such as voluntary disclosure.[1] Neoclassical economic theory assumes individuals are rational optimizers, and is often characterized as providing no role through which individuals could exert idiosyncratic influence on corporate outcomes (e.g., Weintraub 2002; Bertrand and Schoar 2003). Agency theory proposes a less extreme, but still limited role through which individual managers could idiosyncratically influence corporate decisions. In this perspective, managers respond rationally to: (1) the firm's economic environment, and (2) the firm's monitoring mechanisms and managers' contractual incentives. As a result, research has typically focused on "representative" agents because individual managers can be induced to make similar choices through appropriate monitoring and contractual incentives. In this sense, neoclassical and agency theories allow little role for noneconomic manager-specific preferences to affect corporate outcomes. Two streams of research in management lead to a similar conclusion, reasoning that: (1) entrenched norms and cultures in large corporations constrain managers' choices (e.g., Lieberson and O'Connor 1972), and (2) similarity in managers' backgrounds and experiences rising through the corporate ranks further limits heterogeneity (e.g., Hitt and Tyler 1991).

In contrast, Hambrick and Mason's (1984) *upper echelons theory* predicts that managers are not effectively interchangeable. Their theory posits that idiosyncratic differences in managers' experiences are associated with differences in important personal values and cognitive styles such as honesty and tolerance of ambiguity, which *can* lead managers to make different choices, particularly in complex situations lacking clear and calculable solutions. Hambrick and Mason (1984) further argue that such preferences can in turn lead to different organizational outcomes. Thus, we investigate whether individual top managers play a significant incremental role in explaining cross-sectional variation in corporate voluntary disclosure choices, after controlling for known economic determinants of disclosure (and firm-specific effects). In other words, do managers exhibit economically significant individual-specific disclosure styles?

To address this question, we construct a panel data set that tracks top managers over the period 1995–2005. Given that firms are associated with both the key independent variable (managers) and the dependent variables (firms' disclosure choices are sticky over time), it is critical to control for firm-specific effects to avoid a correlated omitted variable problem, and to cleanly isolate manager-specific effects. By isolating manager-specific fixed effects on firms' disclosure characteristics, after controlling for known economic determinants of disclosure, firm fixed effects, and time period fixed effects, we can assess not only the existence, but also the magnitude, of individual managers' incremental effects on their firms' voluntary disclosures. We refer to this *incremental* manager-specific fixed effect as the manager's *style*.

Our results suggest top managers exert economically significant individual-specific influence over five aspects of management forecasts: forecast frequency, forecast precision, news conveyed

[1] Our study focuses on properties of one of the most important and widely investigated forms of voluntary financial disclosure—management earnings forecasts—that Hirst et al. (2008, 315) characterize as "one of the key voluntary disclosure mechanisms by which managers establish or alter market earnings expectations, preempt litigation concerns, and influence their reputation for transparent and accurate reporting."

by the forecast, and the bias in and accuracy of the forecast. We assess the economic significance of the manager-specific fixed effects along three dimensions. First, adding manager-specific fixed effects to the base models (that include the economic determinants, firm-specific effects, and time-period specific effects) increases explanatory power by 7.6 to 10.3 raw percentage points, translating to increases of 29.9 percent to over 400 percent relative to the base models' adjusted R^2s. Second, manager-specific fixed effects are pervasive: between 38 percent and 51 percent of the individual manager-specific fixed effects are significant at the 10 percent level, much greater than expected by chance. Third, the distributions of managers' fixed effects reveal economically material differences. For example, replacing a manager at the 25th percentile of the distribution for forecast frequency with one at the 75th percentile more than triples the number of forecasts.

Having established that individual managers exhibit significant idiosyncratic disclosure styles—contrary to implicit assumptions of traditional financial economics research—our second question is: How do managers' personal experiences influence their own disclosure styles?

Upper echelons theory suggests that observable demographic characteristics of managers' backgrounds reflect key formative experiences that are associated with managers' unique cognitive styles and values. For example, this literature argues that managers' functional career tracks (e.g., marketing, accounting, legal) affect their preferences (e.g., Hambrick and Mason 1984; Jensen and Zajac 2004). We find that, on average, managers promoted from legal backgrounds tend to guide expectations down (reflecting greater sensitivity to litigation risk) and managers promoted from accounting and finance develop more precise disclosure styles that are conservative in underestimating upcoming earnings. Consistent with evidence that individuals who experienced the aftermath of the Great Depression and associated stock market crash are more conservative (e.g., Malmendier and Nagel 2008; Zemke et al. 2000), we find that managers born before World War II are more conservative in being reluctant to forecast. We also find that military experience leads managers to develop styles favoring more precise forecasts that guide expectations down, consistent with lower tolerance for ambiguity, conservatism, and valuing honesty and integrity over self-interest (e.g., Soeters 1997; Goertzel and Hengst 1971; Franke 2001). Finally, recent critics assert that M.B.A. programs overemphasize the pursuit of self-interest (e.g., Gintis and Khurana 2008). We find that M.B.A.s tend to guide expectations upward, but their forecasts are more accurate. Coupled with our evidence on the magnitude and pervasiveness of the manager-specific fixed effects, finding that managers' unique disclosure styles exhibit plausible associations with their distinctive permanent personal demographic characteristics confirms that we are capturing systematic long-lived differences in managers' unique styles (and not just random noise).

Our study makes several contributions. First, it responds to the Brochet et al. (2009, 32) suggestion that "research may explore whether disclosure policies ... vary systematically with managerial style." We provide evidence that—in contrast to the implicit assumption in most prior financial economics research—this hitherto unexplored personal style dimension plays a significant incremental role in explaining cross-sectional differences in voluntary corporate financial disclosure, even after controlling for economic determinants of disclosure and firm-specific (and time-period-specific) fixed effects. Our evidence that not only CEOs, but also CFOs and General Counsels, exhibit distinct disclosure styles responds to Finkelstein and Hambrick's (1996) call for research on the influence of executives beyond the CEO level.

Our second contribution is a response to the Hirst et al. (2008, 316) call for a better understanding of the choices managers make once they decide to issue an earnings forecast (e.g., forecast precision), which Hirst et al. (2008) suggest is an important direction for both theory development and empirical research. Our study provides empirical evidence that managers develop unique and economically significant styles concerning detailed attributes of voluntary disclosure, and these styles reflect their personal demographic characteristics such as functional

career track, age cohort, military experience, and education. Collectively, our evidence suggests that exploring the roles individual managers play in other financial reporting choices is a fruitful avenue for future research.

Third, we contribute more broadly to the nascent manager-effects literature. While Bertrand and Schoar (2003) show that managers matter in core strategic operational and financing decisions, the magnitude of managers' influence varies widely: effects are generally greater for high-profile strategic decisions like acquisitions, and smaller for operational financing decisions like cash holdings. Our evidence that managers not only matter, but matter significantly in seemingly second-order decisions such as the details of firms' voluntary disclosure decisions, extends the range of corporate decisions in which idiosyncratic manager styles are known to play an economically significant role. More importantly, our study contributes beyond existing research by integrating a broad set of literatures to provide conceptual bases for considering: (1) when and why managers might—or might *not*—matter, and (2) why and how characteristics of managers' personal backgrounds likely affect their idiosyncratic disclosure styles.

Finally, from a practical perspective, evidence that disclosures reflect manager-specific styles suggests that initiatives intended to remedy disclosure problems such as upward-biased forecasts (e.g., stock "hyping" observed by Lang and Lundholm [2000]) may be more effective if they target the individual manager, as the Sarbanes-Oxley Act does in holding managers responsible for the firm's financial reports.

II. CONTROVERSY OVER THE ROLE INDIVIDUAL MANAGERS PLAY IN CORPORATE DECISIONS

Differing Perspectives on the Role Individual Top Managers Play in Corporate Decisions

Neoclassical economic theory is often characterized as allowing no role through which idiosyncratic differences among individuals can affect corporate outcomes (e.g., Weintraub 2002). For example, in "a neoclassical view of the firm ... top managers are homogeneous and selfless inputs into the production process ... [and] different managers are regarded as perfect substitutes for one another" (Bertrand and Schoar 2003, 1173). *Agency theory* allows individuals to differ in attributes such as effort aversion, but typically focuses on "representative" agents because monitoring and contractual incentives can induce individuals to make similar choices (e.g., Christensen and Feltham 2003). Grounded in these traditions, archival research has typically abstracted from noneconomic manager-specific effects on corporate choices.[2]

Two streams of research in strategic management similarly conclude that managers are largely interchangeable. First, the *external control perspective* suggests that environmental and organizational constraints such as entrenched norms and cultures limit managers' choices (e.g., Lieberson and O'Connor 1972; Hannan and Freeman 1977). *Legitimacy constraints* are particularly limiting, as managers conform to external expectations of rationality by imitating other managers, increasing homogeneity (e.g., Spender 1989; Hambrick et al. 1993; Chalmers and Godfrey 2004). Second, the *socialization and selection* processes top managers experience rising through the corporate ranks further limits heterogeneity (Hitt and Tyler 1991; Hambrick 2007), with DiMaggio and Powell (1983, 152) asserting "such mechanisms create a pool of almost interchangeable individuals who occupy similar positions across a range of organizations and possess a similarity of orientation and disposition." For example, most CEOs of *Fortune* 1000 firms are white males with college degrees, many from elite institutions (Bhargava and Jespersen 1993; Datz 2000). A sig-

[2] A few studies examine how manager-specific economic incentives such as stock-based compensation and insider trading affect managers' propensity to issue a management forecast (e.g., Noe 1999; Nagar et al. 2003; Cheng and Lo 2006). In contrast, we focus on noneconomic manager-specific effects on attributes of management forecasts.

nificant stream of empirical research in strategic management concludes that idiosyncratic influence of individual top managers is limited (e.g., Lieberson and O'Connor 1972; Salanick and Pfeffer 1977; Hitt and Tyler 1991).

In contrast, Hambrick and Mason's (1984) *upper echelons theory* suggests that individual managers matter. A recipient of the *Academy of Management Review*'s "Frame-Breaking Innovative Theory Award" in 2006 (Kilduff 2006), the study points out that simple inventory and credit policy decisions may have calculable solutions. But in complex ambiguous situations plagued by multiple and often incompatible goals, managers operate within the bounds of rationality, and within these bounds their choices can be influenced by their idiosyncratic experiences and values (Hambrick and Mason 1984; Finkelstein and Hambrick 1996; Hambrick 2007). (Voluntary disclosure is such a complex, ambiguous situation requiring tradeoffs among multiple conflicting goals, such as: (1) protecting proprietary information that competitors could use against the firm versus (2) mitigating litigation risk [e.g., Dye 1986; Evans and Sridhar 2002].) Hambrick and Mason's (1984) theory spurred a stream of empirical research in management, much of which concludes that managers affect firm choices (e.g., Smith and White 1987; Thomas et al. 1991; Tihanyi et al. 2000; Naranjo-Gil and Hartmann 2006; Crossland and Hambrick 2007). However, *this research does not isolate manager effects after controlling for firm effects.*[3] If unmodeled firm characteristics affect corporate choices, then by confounding managers' effects with firm effects, this literature likely overstates any "manager" effects.

Archival research in finance and economics is just starting to explore whether individual managers impose idiosyncratic influence on corporate decisions. The seminal study in this genre, Bertrand and Schoar (2003), develops an innovative design that disentangles the manager effect from the firm effect by tracking managers across firms over time. They find that managers develop unique individual-specific styles in operational and financing decisions. Individual managers' incremental effects are generally greater for key strategic decisions like acquisitions, but smaller for operational financing policies such as cash holdings. If managers have any idiosyncratic influence on corporate decisions, then it is most likely to be on high-profile core strategic decisions where top managers are most interested and actively involved.

Ex ante, it is unclear whether managers would exert economically significant idiosyncratic influence on their firms' *voluntary financial disclosure choices.* Individual top managers' effects on voluntary disclosure hinge on: (1) the extent to which managers act as economic agents versus individuals with distinct communication styles, and (2) the discretion managers enjoy over the strategic choice (Hambrick 2007). Top managers who act primarily as economic agents are likely more interested in and involved in high-profile strategic operational and financing decisions (e.g., acquisitions, strategic marketing, production expansion), but are less interested in and less involved in seemingly secondary decisions about detailed attributes of management earnings forecasts. This suggests that top managers would have limited (if any) effect on the firm's voluntary disclosure choices. On the other hand, research in psychology concludes that humans develop long-lived, individual-specific communication styles (e.g., Pennebaker and King 1999). If top managers have distinct communication styles, such as valuing openness and transparency, then this would likely affect the characteristics of the firm's voluntary disclosures. With respect to discretion, top managers can *in principle* exercise discretion over management earnings forecasts, but their discretion is limited by the *ex post* verifiability of the disclosures and stickiness of firms' disclosure policies over time (Healy et al. 1999; Lang 1999).

[3] The only exception we found that does control for firm fixed effects is Crossland and Hambrick's (2007) study of the cross-cultural differences in the effect of CEOs on measures of firm performance (e.g., return on assets).

To summarize, certain theories and evidence suggest managers will have at most limited effect on their firms' voluntary disclosure choices, whereas other theories and evidence suggest the opposite. It is therefore an empirical question whether top managers exert economically significant individual-specific influence over their firms' voluntary disclosures. Thus, we investigate whether individual managers develop unique disclosure styles, paying special attention to the magnitude of any such manager-specific effects.[4]

What Demographic Characteristics of Managers' Personal Backgrounds Might Influence Their Disclosure Styles?

To the extent that managers' backgrounds, socialization, and selection processes create similar experiences and induce similar values and cognitive styles, the demographic characteristics of their personal backgrounds are unlikely to explain any differences in their disclosure styles (e.g., DiMaggio and Powell 1983; Hitt and Tyler 1991; Hambrick 2007). On the other hand, upper echelons theory suggests cross-sectional differences in managers' demographic characteristics, such as their functional career track, age, and education, differentially shape managers' values and cognitive bases, which in turn can affect their managerial styles (e.g., Hambrick and Mason 1984; Hitt and Tyler 1991). We focus on the role of observable demographic characteristics of managers' backgrounds since upper echelons theory recommends that "An emphasis on background characteristics, rather than psychological dimensions, seems essential" (Hambrick and Mason 1984, 196; see also Hambrick 2007, 335).

Whereas certain demographic background characteristics are observable and reliably measurable, top managers (especially of large corporations) are reluctant to submit to batteries of psychological tests, which are "too general and detached from executive issues to be useful for studying top managers" (Finkelstein and Hambrick 1996, 46). That is, psychological test instruments intended for the general population are not tailored to identify finer psychological traits that might distinguish among successful professional executives. It is therefore not surprising that research using psychological measures has not found significant effects. For example, Sorter et al. (1964) and Faircloth and Ricchiute (1981) find that personality characteristics do not explain accountants' preferred reporting alternatives. In contrast, Faircloth and Ricchiute (1981, 53) find that objective personal demographic characteristics such as education do play a significant role, prompting them to recommend that "additional research should pay more attention to the effects of demographic information, which may in fact be more informative than many previously tested variables such as personality characteristics."

To guide our exploration of the relation between managers' demographic characteristics and their unique disclosure styles, we conducted a comprehensive review of a broad set of literatures ranging from strategic management, to career counseling, to sociology, to psychology, to military science, and to business education. We searched for theories or evidence suggesting how demo-

[4] A recent working paper by Brochet et al. (2009) investigates the effect of CEO or CFO turnover on the frequency and precision of management forecasts. When a new CFO is appointed, on average firms that provided frequent guidance before the turnover experience a temporary decrease in the quantity and precision of management earnings forecasts. When such firms appoint a new CEO, the decline in the frequency of management forecasts tends to be longer-lived (documented up to two years), but there is no effect on forecast precision. Thus, Brochet et al. (2009) document the effect of a managerial transition *per se* on the firm's disclosure policy. Our study is fundamentally different. Rather than focusing on the short- or medium-term effects of a shock resulting from a managerial transition *per se*, we focus on whether a given individual manager exhibits a persistent disclosure style measured over at least six years across multiple employers.

Another working paper by Ge et al. (2009) explores the effects of CFOs on choices that affect mandatory financial reporting. They conclude that CFOs exert manager-specific fixed effects on GAAP earnings-related choices indicative of earnings management (e.g., abnormal accruals). They do not explore the effects of individual managers on *voluntary management earnings forecasts*, which is the focus of our study.

graphic characteristics highlighted in Hambrick and Mason's (1984) upper echelons theory might plausibly be associated with cross-sectional differences in managers' unique disclosure styles. Our review suggests four demographic characteristics of managers' personal backgrounds—functional career track, age cohort, other career experience (specifically, military experience), and M.B.A. education—likely influence managers' unique disclosure styles.[5]

Functional Career Track

Hambrick and Mason's (1984, 200) upper echelons theory suggests a manager's primary functional career track affects his/her choices because "career experiences partially shape the lenses through which they view current strategic opportunities and problems." Socialized in their functional area's mindset, managers tend to converge on solutions to common business problems (Finkelstein and Hambrick 1996, 93). Empirical research confirms that managers pursue strategies in line with their own functional expertise (e.g., Smith and White 1987; Thomas et al. 1991; Jensen and Zajac 2004). Prior research characterizes individuals in technical financial functions as conventional, orderly, and inhibited (Holland 1973; 1997), suggesting they may adopt conservative disclosure styles. (We consider fewer forecasts, downward guidance, or guidance that undershoots upcoming earnings as various dimensions of a conservative disclosure style.) Managers from technical functions are also less tolerant of ambiguity (Holland 1966; 1997, 27), and Hambrick and Mason (1984) posit that managers from finance favor more budget detail and thoroughness, which suggests managers from finance or accounting may develop more precise communication styles. Executives with legal backgrounds are more sensitive to litigation risk, and so likely favor communication styles that do not promise too much.[6]

Age Cohort

Upper echelons theory suggests that the manager's age cohort can affect his/her values, cognitive styles, and thus his/her decisions (Hambrick and Mason 1984). Schuman and Scott (1989) show that the events an age cohort (individuals born in the same era) experiences during its youth affect later attitudes. Zemke et al. (2000) classify managers born before World War II as an important cohort, shaped by the lingering sacrifices and shortages of the Great Depression, the Dust Bowl that wiped out the Great Plains farms, and the war years, all of which inculcated conservative values. Emerging research in financial economics finds that cohorts who experienced lower stock returns during their investing lives are more conservative (Malmendier and Nagel 2008). Blanchard (1993) ascribes the decline in the equity risk premium to receding memories of the crash and the Great Depression, and Palsson (1996) finds that age is associated with greater risk aversion in portfolio holdings. Finally, Bertrand and Schoar (2003) find that managers from older birth cohorts choose lower levels of corporate expenditures, lower leverage, and larger cash holdings, consistent with conservatism. These arguments suggest executives born before World War II may develop more conservative communication styles, such as being less likely to issue forward-looking statements for fear such disclosures may prove inaccurate, *ex post.*

---

[5] Upper echelons theory suggests that certain demographic characteristics affect managers' cognitive styles and values, but we had to search elsewhere to support how these characteristics might plausibly affect managers' *disclosure* styles. As a result of this search, we explored two additional demographic characteristics beyond upper echelons theory: gender and whether the manager is from the firm's founding family. Unfortunately, members of founding families of S&P 1500 firms do not move to similar positions at other S&P 1500 firms. Consistent with the literature documenting the homogeneity of managers' backgrounds, our sample includes only six female managers.

[6] Bagley (2008, 379) argues that "legally astute" top managers make different choices because they "appreciate the importance of meeting society's expectations of appropriate behavior … and of treating stockholders fairly."

Military Experience

Upper echelons theory suggests that an executive's other career experiences can affect his/her decisions. We posit that military service is a different form of career experience that likely affects a manager's disclosure style. Military personnel tend to be less tolerant of ambiguity/uncertainty (e.g., Goertzel and Hengst 1971; Soeters 1997), suggesting they may favor more precise communication styles. A common theme is "service before self" (Franke 2001), as the Army describes military values as "duty, selfless service, courage, integrity … and a sense of justice" (Headquarters, Department of the Army 1993, 1–2). Individuals choosing military service tend to be conservative (e.g., Goertzel and Hengst 1971; Franke 2001). They are likely to agree that "honesty is the best policy," that "one should take action only when one is sure it is morally right," and are less likely to believe it is acceptable to "cut corners" (Franke 1998; 2001). They place relatively lower value on income and promotion opportunities (Soeters 1997). Such values suggest managers with military experience may favor styles that tend toward prompt and unbiased disclosure of unfavorable information (i.e., honesty about bad news).

M.B.A. Education

Upper echelons theory predicts managers holding M.B.A. degrees develop different styles than those without such educational backgrounds. Finkelstein and Hambrick (1996, 104) suggest M.B.A.s are part of the social and business elite who value conformity and conventionality. M.B.A.s are also likely aware of the penalties the market exacts when firms fail to meet their own earnings forecasts (Chen 2004). These arguments suggest that M.B.A.s may adopt more conservative disclosure styles. On the other hand, post-Enron era critics have asserted M.B.A. programs overemphasize neoclassical economics, where rational managers pursue their own self-interest (e.g., Ghoshal 2005; Gintis and Khurana 2008).[7] Self-reported cheating levels are higher in M.B.A. than other graduate programs (McCabe et al. 2006), and individuals who behave dishonestly as students are more likely to behave dishonestly in the workplace (Nonis and Swift 2001), suggesting M.B.A.s might be more likely to issue disclosures that spin the firm's outlook.

III. RESEARCH DESIGN

The Existence and Magnitude of Individual Top Managers' Voluntary Disclosure Styles

Sample and Data

Firms' disclosure choices are sticky over time (Healy et al. 1999; Lang 1999), but as Hirst et al. (2008) point out, empirical models leave most of the cross-sectional variation in voluntary disclosure choices unexplained. This suggests that unidentified firm characteristics likely affect firms' disclosure choices. Because firms are associated with both the key independent variable (managers) and the dependent variables (disclosure choices), it is necessary to control for firm-specific fixed effects to avoid misattributing firm effects to managers.

We therefore use a research design with strong controls for firm effects. We adopt Bertrand and Schoar's (2003) longitudinal design that tracks managers over time, and requires managers to work for two or more employing firms, with at least three years at each employer so the manager has time to "imprint" his/her style. (If the manager works at only one firm, then his/her effect is confounded with the firm effect and it is not possible to cleanly separate the manager effect from the firm effect.) Figure 1 illustrates this design for two firms. For each firm, we control for firm-specific fixed effects measured over the entire sample period. To isolate the effect of an

[7] Bertrand and Schoar (2003) provide some evidence that managers with M.B.A.s make more aggressive resource allocation decisions. Because M.B.A. education focuses on operational and financing decisions and less on disclosure, evidence that an M.B.A. degree is associated with manager-specific effects on operating and financing decisions does not necessarily imply that there will be similar effects on managers' voluntary disclosure decisions.

FIGURE 1
Research Design for Two Illustrative Firms

	FIRM 1	FIRM 2
Period 1 (at least 3 consecutive years)	***Manager A***	Manager C
Period 2 (at least 3 consecutive years)	Manager B	***Manager A***

Manager A's fixed effect is measured over Firms 1 and 2, after controlling for: (1) the fixed effect of Firm 1 (estimated over the entire sample period), (2) the fixed effect of Firm 2 (estimated over the entire sample period), (3) fixed effects for each year, and (4) all of the time-varying economic determinants of disclosure.

individual manager on disclosure—*incremental to the effects of the firms*—we measure the effect of Manager A across the two employing firms, after controlling for Firm 1's fixed effect, Firm 2's fixed effect, year fixed effects, and well-documented economic determinants of voluntary disclosure.[8] Manager A's style is measured by his/her fixed effect.

Manager A's fixed effect reflects how Firm 1's forecast attributes in period 1 deviate from Firm 1's average forecast attributes, and how Firm 2's forecast attributes in period 2 deviate from Firm 2's average forecast attributes (after also abstracting from the effects of known economic determinants and year-specific effects). Thus, our estimated manager-specific fixed effects capture the incremental effects of a manager at multiple employers, *after controlling for any firm-specific effects*, over six (or more) years. This method also reduces concerns about correlated omitted variables, which to explain our results would have to vary over time and across firms in the same pattern as the movement over time and across firms of managers with unique disclosure styles.

From Execucomp, we record the five highest-paid managers for S&P 1500 firms. Our sample period starts in 1995 when First Call's Company Issued Guideline (CIG) database became a more complete source of management forecasts (Anilowski et al. 2007).[9] From 1995 to 2005, we identify 303 top managers that work for at least two different firms and stay employed by each firm for at least three years. We focus on managers whose last position is CEO, CFO, or General Counsel because these positions are more likely to influence voluntary disclosure. We match each firm-year with management forecasts reported in CIG.

Model and Variable Definitions

We investigate five attributes of management earnings forecasts: frequency, precision, news, direction of bias relative to actual earnings, and accuracy. The frequency of managers' forecasts (*F_Freq*) is the number of earnings-related forecasts the firm issues during the year (which is 0 if

[8] Following Bertrand and Schoar (2003), we retain all firm-years with data only in the estimation of the firm fixed effects (i.e., including Manager B observations for estimation of Firm 1's fixed effect and Manager C observations for estimation of Firm 2's fixed effect).

[9] Ajinkya et al. (2005, 350) conclude that CIG is a comprehensive source of management forecast information, after performing two small-sample tests (in 1997 and in 2000) matching CIG forecasts with those identified in a keyword search of Factiva (formerly the Dow Jones News Retrieval Service). Choi and Ziebart (2004) report similar results for the period 1993–1996.

the firm issues no forecasts).[10] We code forecast precision (*F_Precision*) as 3, 2, 1, and 0, for point, range, open-ended, and general impression forecasts (e.g., Bamber and Cheon 1998; Baginski et al. 2002).[11] We measure forecast news as the difference between the forecast EPS (or midpoint of the range forecast) and the prevailing median analyst forecast, scaled by lagged stock price (*F_Goodnews*). We measure forecast bias (*F_BiasUP*) by comparing the management EPS forecast (using point forecasts and the midpoint of range forecasts) to *ex post* realized EPS.[12] Forecasts that exceed actual EPS are coded as +1 (i.e., optimistic forecasts), those that equal actual EPS are coded as 0, and those that fall short of actual EPS are coded as −1 (i.e., pessimistic forecasts).[13] Finally, we measure the inverse of forecast accuracy (*F_AbsError*) as the absolute value of the difference between the management EPS forecast (using point forecasts and the midpoint of range forecasts) and actual EPS, scaled by lagged stock price, following Ajinkya et al. (2005). Exhibit 1 summarizes our variable definitions.

To test whether individual managers play unique, individual-specific roles in firms' voluntary financial disclosure, we estimate the following models:

$$Y_{it} = \alpha + \beta X_{it} + \tau_t + \gamma_i + \varepsilon_{it}; \tag{1}$$

$$Y_{it} = \alpha + \beta X_{it} + \lambda_{Managers} + \tau_t + \gamma_i + \varepsilon_{it}; \tag{2}$$

$$Y_{it} = \alpha + \beta X_{it} + \lambda_{CEOs} + \tau_t + \gamma_i + \varepsilon_{it}; \tag{3}$$

$$Y_{it} = \alpha + \beta X_{it} + \lambda_{CFOs} + \tau_t + \gamma_i + \varepsilon_{it}; \tag{4}$$

$$Y_{it} = \alpha + \beta X_{it} + \lambda_{Counsels} + \tau_t + \gamma_i + \varepsilon_{it}. \tag{5}$$

Our unit of analysis is a firm-year. Y_{it} is the management forecast attribute for firm i in year t. When firms issue multiple forecasts in a year, we use firm-year averages for precision, news, bias, and accuracy.[14] The vector X controls for a comprehensive set of economic determinants of voluntary disclosure documented in prior literature, as summarized in Figure 2 and explained in the Appendix. $\lambda_{Managers}$, our main variables of interest, are estimated coefficients for indicator variables corresponding to each individual manager, and thus represent the individual managers' fixed

[10] *F_Freq* is skewed, but our inferences are robust if we redefine *F_Freq* as the natural log of 1 plus the number of forecasts. Our inferences are also robust when we conduct the analysis using a fixed effect count model (Poisson model), per Cameron and Trivedi (2005).

[11] Our main analysis includes forecasts of all specificity levels. However, we obtain similar inferences using an alternative (inverse) measure of specificity where we code point forecasts as zero and range forecasts as the magnitude of the range, scaled by lagged share price.

[12] We compare management earnings forecasts listed on First Call to actual earnings as reported by First Call to ensure consistency, following Ajinkya et al. (2005).

[13] Our main analysis focuses on the direction of the bias, because the forecast error variable captures the distance of the forecast from *ex post* actual EPS. Nonetheless, we obtain similar inferences if we define *F_BiasUP* as the difference between the management forecast and *ex post* actual EPS, scaled by lagged share price.

[14] We conduct our analysis at the firm-year level because our goal is to assess the extent to which individual managers affect overall firm disclosure policy. Untabulated sensitivity tests reveal that our results continue to hold if we use the last forecast issued in a year. Any dependence among multiple observations from the same firm does not affect our conclusions, because they continue to hold when we allow clustering of the error term at the firm level.

EXHIBIT 1

Variable Definitions

Dependent Variables

F_Freq_{it} = number of management earnings forecasts firm i issued in year t;

$F\_Precision_{it}$ = average precision of management earnings forecasts firm i issued in year t, where point forecasts are coded as 3, range forecasts are coded as 2, open-ended forecasts are coded as 1, and qualitative forecasts are coded as 0;

$F\_Goodnews_{it}$ = average news conveyed by forecasts firm i issued in year t, measured as the difference between the forecast EPS (or midpoint of the range forecast) and the prevailing median analyst forecast, scaled by lagged stock price;

F_BiasUP_{it} = average bias (relative to *ex post* actual earnings) in forecasts firm i issued in year t.[a] Forecasts that exceed actual earnings are coded as $+1$, forecasts that equal actual earnings are coded as 0, and forecasts that fall short of actual earnings are coded as -1 (for range forecasts, we code this variable based on the midpoint of the range); and

$F\_AbsError_{it}$ = inverse of forecast accuracy, measured as the average absolute error in the forecasts that firm i issued in year t, relative to *ex post* actual earnings. Specifically, we use the absolute value of the difference between the management forecast (using point forecasts and the midpoint of the range forecasts) and actual EPS, scaled by lagged stock price, multiplied by 100 to yield an error in percentage terms.

Control Variables

$|\Delta EPS_{it}|$ = absolute value of the change in firm i's earnings per share from year $t-1$ to t, deflated by stock price at the end of year $t-1$;

$Disp_{it}$ = standard deviation of analysts' forecasts of firm i's year t earnings, divided by the absolute value of the median forecast;

EPS_UP_{it} = indicator variable coded as 1 if firm i's EPS in year t is greater than or equal to its EPS in year $t-1$, and 0 otherwise;

$Loss_{it}$ = indicator variable that equals 1 if firm i reports a loss in year t (Compustat item 172, net income), and 0 otherwise;

$R\&D_{it}$ = firm i's expenditures on research and development in year t, scaled by its total assets;

$Mkt-Bk_{it}$ = market value of firm i's common equity divided by the book value of its common equity, at the end of year t;

$Outside\_Dir_{it}$ = percentage of members of firm i's board of directors during year t, who are not also officers of the firm, from the Investor Responsibility Research Center (IRRC);

$\%Inst_{it}$ = percentage of firm i's common stock held by institutions in year t, from Thomson Financial;

$Lsize_{it}$ = natural log of the market value of firm i's common equity in year t;

$\#Analysts_{it}$ = number of analysts providing earnings forecasts for firm i in year t;

$Litigation_{it}$ = coded as 1 if the firm is a member of one of the following high-litigation-risk industries: SIC codes 2833–2836 (biotechnology), 3570–3577 and 7370–7374 (computers), 3600–3674 (electronics), 5200–5961 (retailing), and 8731–8734 (R&D service), and suffers a 20 percent or greater decrease in earnings; and 0 otherwise;

$ClassAction_{it}$ = 1 if the firm is involved in a securities lawsuit in year t, per the Stanford Security Class Action Clearinghouse;

$Restructuring_{it}$ = 1 if the firm was engaged in a restructuring during year t;

$Restatement_{it}$ = 1 if the firm had a restatement in year t, per the General Accounting Office's Accounting Restatement database;

$Acquisition_{it}$ = 1 if the firm has a merger or acquisition in year t, per footnote 1 in the annual Compustat file; and

$F\_Horizon_{it}$ = number of days between the forecast date and the end of the fiscal period of the forecasted earnings number. We scale this forecast horizon by 360 for annual forecasts or by 90 for quarterly forecasts, so the variable measures the proportion of the fiscal period remaining at the date firm i issues the forecast.

(continued on next page)

EXHIBIT 1 (continued)

Additional Firm Characteristics

Mve_{it} = market value of firm i's common equity at the end of year t, in millions;
$Sales_{it}$ = sales revenue of firm i in year t, in millions;
ROA_{it} = firm i's net income in year t divided by lagged total assets;
Lev_{it} = firm i's long-term debt (Compustat item 9 + item 34) divided by (long-term debt plus the book value of common equity (Compustat item 60)); and
$Beta_{it}$ = firm i's equity beta for the fiscal period, calculated using daily equity returns and a single factor CAPM.

Manager Demographic Characteristics

Acct/Fin = 1 if managers rise from accounting/finance backgrounds, and 0 otherwise;
Legal = 1 if managers rise from legal backgrounds, and 0 otherwise;
PreWWII = 1 if managers were born before World War II, and 0 otherwise;
Military = 1 if managers have military experience, and 0 otherwise; and
MBA = 1 if managers hold an M.B.A. degree, and 0 otherwise.

[a] All actual EPS numbers are before extraordinary items.

FIGURE 2
Detailed Model Specification

Dependent Variable (Y_{it})	Vector of Control Variables (X_{it})
Number of forecasts (F_Freq_{it})	$\|\Delta EPS_{it}\|$, $Disp_{it}$, EPS_UP_{it}, $Loss_{it}$, $R\&D_{it}$, $Mkt\text{-}Bk_{it}$, $Outside\_Dir_{it}$, $\%Inst_{it}$, $LSize_{it}$, $\#Analysts_{it}$, $Litigation_{it}$, $ClassAction_{it}$, $Restructuring_{it}$, $Restatement_{it}$, $Acquisition_{it}$
Forecast precision ($F\_Precision_{it}$) Forecast news ($F\_Goodnews_{it}$) Forecast bias (F_BiasUP_{it}) Forecast accuracy ($F\_AbsError_{it}$)	$\|\Delta EPS_{it}\|$, $Disp_{it}$, EPS_UP_{it}, $Loss_{it}$, $R\&D_{it}$, $Mkt\text{-}Bk_{it}$, $Outside\_Dir_{it}$, $\%Inst_{it}$, $LSize_{it}$, $\#Analysts_{it}$, $Litigation_{it}$, $ClassAction_{it}$, $Restructuring_{it}$, $Restatement_{it}$, $Acquisition_{it}$, $F\_Horizon_{it}$

effects. τ_t controls for year-specific fixed effects, and γ_i controls for firm-specific fixed effects.[15]

[15] We estimate models 1–5 using a firm fixed effect model (XTREG in STATA) that estimates coefficients for the economic determinants, manager indicator variables, and year indicator variables. The estimation eliminates the unobserved time-invariant firm characteristics in models 1–5 (i.e., the γ_i) by a within-transformation that demeans each variable for each firm (Wooldridge 2002, 267). An alternative is OLS estimation with explicit firm indicator variables. That approach generates the same coefficient estimates on the economic determinants (the vector X) and the manager ($\lambda_{Managers}$) and time period effects (τ_t), but it is computationally intensive, infeasible for models with thousands of firms (as is the case in one of our sensitivity tests), and it also generates higher adjusted R^2. So our reported adjusted R^2s are more conservative.

Model 1 includes known economic determinants of voluntary disclosure choices as well as time- and firm-specific fixed effects. Model 2 adds the manager-specific fixed effects. This enables us to investigate not only the existence of, but also the *magnitude* of, individual managers' *incremental* effects on their firms' voluntary disclosures after controlling for the known economic determinants, as well as time- and firm-specific fixed effects. Models 3, 4, and 5 test for effects of CEOs, CFOs, and General Counsels, respectively.

Testing Whether Demographic Characteristics of Top Managers' Personal Backgrounds Are Associated with Their Unique Disclosure Styles

Data, Model, and Variable Definitions

We hand-collect demographic characteristics of managers' backgrounds from Marquis *Who's Who*, Zoominfo, Mergent Online, *Hoover's*, NNBD.com, and company websites. To test whether these characteristics contribute to managers' individual disclosure styles, we regress the fixed-effect coefficient estimates from Model 2 (after controlling for economic determinants, time- and firm-specific fixed effects), on managers' personal demographic characteristics:

$$\lambda_{m,Yi} = \alpha_0 + \gamma_1 Acct/Fin + \gamma_2 Legal + \gamma_3 PreWWII + \gamma_4 Military + \gamma_5 MBA + \varepsilon. \quad (6)$$

$\lambda_{m,Yi}$ are the manager-specific fixed effects estimated in Model 2, where m represents the manager and Y represents the disclosure attribute. *Acct/Fin* distinguishes managers rising from accounting or finance backgrounds, and *Legal* distinguishes managers with legal backgrounds.[16] Following Zemke et al. (2000), we classify managers into two age cohorts—managers born before World War II (*PreWWII)* versus those born after World War II.[17] *Military* indicates the manager has military experience, and *MBA* indicates the manager earned an M.B.A. degree.

IV. RESULTS

Descriptive Statistics

Table 1 compares our sample firm-years to the population with at least three consecutive years of data in Execucomp that also have at least one analyst following the firm, as reported by First Call. Our sample firms tend to be larger, because managers of larger firms are more likely to move to another firm in the S&P 1500 population covered by Execucomp. This biases *against* finding a manager-specific effect, because managerial discretion—and, hence, any manager-specific effect—declines with company size (Finkelstein and Hambrick 1996; Bertrand and Schoar 2003). Our sample firms are more levered, but similar in profitability and risk (beta).

On average, our sample firms issue 1.77 forecasts each year, slightly more than the population average of 1.48 ($p \leq 0.05$). Our sample firms' forecasts are less biased (relative to *ex post* actual earnings) than those of the population ($p \leq 0.05$). However, forecasts issued by our sample and the population exhibit similar precision, news (relative to prevailing analysts' forecasts), and accuracy. The average forecast gives a range of values, guides expectations downward, and is downward-biased relative to actual earnings.

Turning to the economic determinants, our sample firms are slightly less likely to enjoy an increase in earnings during the forecast year, and have a larger proportion of outside directors,

[16] We expect the accounting/finance and legal career tracks to be most relevant to voluntary earnings-related disclosure choice. Separate analysis of executives promoted from production and marketing tracks does not affect our inferences, nor were the coefficients on the marketing or production career track variables significant.

[17] Zemke et al. (2000) identify those born before 1943 as the *Veterans* cohort, those born between 1943 to 1960 as the *Baby Boomers*, and those born between 1960 to 1980 as *Generation Xers*. Those born from 1943 to 1946 are often referred to as the *Sandwich* generation, but Zemke et al. (2000) consider them part of the *Baby Boomers* cultural cohort. We combine the Baby Boomers and Generation Xers because our sample includes only a handful of Generation Xers, and separating them does not yield any additional insights.

TABLE 1

Comparison of Sample to the Population in the Intersection of Execucomp and First Call[a]

	Sample Firms				Execucomp/First Call Population			
Variable	**n**	**Mean**	**Std. Dev.**	**Median**	**n**	**Mean**	**Std. Dev.**	**Median**
Firm Characteristics								
*Mve**	3,920	11,066	30,774	2,592	18,987	6,349	21,305	1,282
*Sales**	3,920	7,847	19,578	2,217	18,987	4,281	12,474	1,058
ROA	3,920	0.04	0.16	0.04	18,987	0.04	0.37	0.05
*Lev**	3,909	0.40	0.51	0.40	18,915	0.35	0.98	0.35
Beta	3,915	0.98	0.56	0.87	18,937	0.97	0.56	0.87
Management Forecast Characteristics								
*F_Freq**	3,920	1.77	2.27	1	18,987	1.48	2.11	0
F_Precision	2,169	1.97	0.72	2	9,321	1.96	0.73	2
F_Goodnews	1,920	−0.16	0.58	−0.04	8,193	−0.18	0.69	−0.05
*F_BiasUP**	1,914	−0.12	0.80	−0.11	8,151	−0.20	0.79	−0.33
F_AbsError	1,914	0.58	1.15	0.21	8,151	0.55	1.04	0.20
Economic Explanatory Variables								
|ΔEPS|	3,920	0.07	0.15	0.02	18,987	0.06	0.14	0.02
Disp	3,920	0.11	0.28	0.03	18,360	0.10	0.26	0.03
*EPS_UP**	3,920	0.63	0.48	1	18,987	0.65	0.48	1
Loss	3,920	0.18	0.39	0	18,987	0.17	0.37	0
R&D	3,920	0.03	0.05	0	18,987	0.03	0.06	0
Mkt-Bk	3,920	3.30	3.58	2.34	18,987	3.22	3.35	2.30
*Outside_Dir**	3,920	70.39	16.47	75	18,888	66.07	18.33	69.23
*%Inst**	3,920	63.61	18.99	65.58	18,961	61.34	20.81	63.07
*LSize**	3,920	7.92	1.60	7.86	18,987	7.30	1.57	7.16
*#Analysts**	3,920	14.71	8.94	13	18,987	11.46	8	9
Litigation	3,920	0.10	0.30	0	18,987	0.10	0.30	0
*ClassAction**	3,920	0.04	0.20	0	18,987	0.03	0.17	0
*Restructuring**	3,920	0.35	0.48	0	18,987	0.28	0.45	0
*Restatement**	3,920	0.05	0.21	0	18,987	0.03	0.18	0
Acquisition	3,920	0.21	0.41	0	18,987	0.21	0.40	0
*F_Horizon**	2,169	0.44	0.31	0.47	9,321	0.42	0.33	0.45

* Indicates that our sample and the Execucomp/First Call population differ significantly both on t-tests and nonparametric Wilcoxon rank sum tests ($p < 0.05$).

[a] We compare our sample firms to the population of firms with at least three consecutive years of data in Execucomp, and with at least one analyst following from First Call.

Exhibit 1 defines all variables.

higher institutional ownership, and larger analyst following. Our sample firms also have slightly higher incidences of class action lawsuits, restructurings, and restatements, and their management earnings forecasts have somewhat longer horizon. However, the sample and population exhibit similar absolute changes in EPS, dispersion in analysts' forecasts, incidence of losses, R&D intensity, market-to-book ratios, litigation risk, and acquisition activity. In sum, our sample firms are somewhat larger than the population, and exhibit other differences associated with larger firms, but our sample does not differ on other fundamental economic characteristics such as risk (beta), return on assets, absolute changes in EPS, and market-to-book ratios.

The Existence and Magnitude of Individual Top Managers' Voluntary Disclosure Styles

Main Results

The results of the manager fixed-effects analyses appear in Tables 2–6. The primary inference is that top managers have statistically *and* economically significant fixed effects on their firms' voluntary disclosure choices. We discuss these results in more detail.

Table 2 presents the results for the frequency of forecasts. The F-tests from Model 2 indicate

TABLE 2

Testing Individual Top Managers' Fixed Effects on the Number of Management Earnings Forecasts

F_Freq (n = 3,920)[a]	Economic Determinants, Year- and Firm-Specific Fixed Effects (Model 1)	Economic Determinants *and* Manager Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 2)	Economic Determinants *and* CEOs' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 3)	Economic Determinants *and* CFOs' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 4)	Economic Determinants *and* General Counsels' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 5)
Testing Economic Determinants = 0					
F-statistics	6.77	4.32	5.16	5.38	7.11
(p-value)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Constraints	15	15	15	15	15
Testing Manager Fixed Effects = 0					
F-statistics		7.46	7.10	5.92	5.01
(p-value)		(<0.001)	(<0.001)	(<0.001)	(<0.001)
Constraints		292	177	93	22
Adjusted R^2	33.5%	43.5%	40.2%	36.4%	34.0%
Improvement in Adjusted R^2 Relative to Model (1):					
In Raw Percentage Points		10.0	6.7	2.9	0.5
As a % of Model 1's R^2		29.9%	20.0%	8.7%	1.5%

Distribution of Manager Fixed Effects[b] (n = 292 Managers)

	Manager Fixed Effect
% Significant at 10% Level	51%
Mean Effect	0.00
Median Effect	0.01
25th Percentile	−0.97
75th Percentile	0.96
Difference between 75th and 25th Percentile	1.93

[a] We report the results of five regressions after controlling for economic determinants, and year- and firm-specific fixed effects. We control for heteroscedasticity using White's standard errors. The adjusted R^2 is calculated based on the Within R^2 from XTREG, FE in STATA. Exhibit 1 defines the economic determinants.

[b] The manager fixed effects are the estimated coefficients from Model 2 that controls for economic determinants as well as year- and firm-specific fixed effects in estimating the *incremental* manager fixed effects.

that the manager-specific fixed effects are jointly different from zero, after controlling for economic determinants as well as firm- and time-specific fixed effects ($p < 0.001$). We next assess the economic significance of these manager-specific fixed effects in three ways. We show: (1) the increase in adjusted R^2 from adding the manager fixed effects, both in terms of raw percentage points, and in percentage changes (i.e., the increase in adjusted R^2 as a percentage of the adjusted R^2 from Model 1), (2) that the proportion of statistically significant manager fixed effects is much higher than expected by chance, and (3) that the interquartile range indicates meaningful cross-sectional differences in managers' disclosure styles.

First, even after controlling for known economic determinants of disclosure (as well as time- and firm-specific fixed effects), adding manager fixed effects to the base Model 1 significantly improves explanatory power. Adjusted R^2 increases from 33.5 percent in the base model to 43.5 percent, an increase of ten percentage points in absolute terms and 29.9 percent (10/33.5) in relative terms.[18] Models 3 to 5 show how managers in the CEO, CFO, and General Counsel positions contribute to the increase in explanatory power observed in Model 2.[19]

The second chart in Table 2 displays the two additional indicators of the economic significance of the manager-specific fixed effects estimated in Model 2. Even after controlling for economic determinants, 51 percent of the individual manager-specific effects are significant at the 10 percent level, much higher than expected by chance. The interquartile range shows that replacing a manager at the 25th percentile with one at the 75th percentile increases the number of forecasts issued each year by 1.93, which is material given that the mean number of forecasts our sample firms issue each year is 1.77 (Table 1). Clearly, individual managers' unique fixed effects play an economically significant role in firms' forecasting frequency.

Tables 3–6 show that the economic determinants explain much less of the cross-sectional variation in forecast precision, news, bias, and error (1.8 percent to 12.5 percent) than of forecast frequency (33.5 percent, Table 2), consistent with the Hirst et al. (2008) observation that we know less about the attributes of management forecasts than we do about the decision to issue a forecast. For forecast precision, adding the manager-specific effects to the base Model 1 increases adjusted R^2 from 1.8 percent to 10.0 percent, an increase of 8.2 percentage points in absolute terms (or an increase of 456 percent in relative terms). Similarly, adding the manager-specific fixed effects to the forecast news, bias, and error models increases adjusted R^2 by 9.5, 7.6, and 10.3 percentage points (which translate to increases of 82 percent, 61 percent, and 82 percent in relative terms), relative to Model 1. The increases in adjusted R^2 observed in Models 3 to 5 suggest that CEOs' and CFOs' styles play a significant role in these forecast attributes, but the General Counsels' role is not economically meaningful except for the forecast error attribute.

The second chart in Tables 3–6 displays the distribution of manager-specific fixed effects estimated in Model 2. Between 38 percent and 46 percent of the individual manager fixed effects are significant at the 10 percent level, much higher than expected by chance. The interquartile ranges further support the materiality of the individual manager effects. Table 3 shows that replacing a manager at the 25th percentile with one at the 75th percentile increases average precision by 0.75, which is material relative to the unconditional mean of 1.97 (per Table 1). Turning to

[18] The increase in adjusted R^2 from adding manager-specific fixed effects compares favorably with the increase in adjusted R^2 from adding manager-specific fixed effects to models of other corporate decisions. For example, Bertrand and Schoar (2003) find that adding manager-specific fixed effects to models of investment and financial policies increases adjusted R^2 by an average of five percentage points across eight decisions.

[19] Following Bertrand and Schoar (2003), our primary analyses are based on the manager's position at the last firm. When we require managers to retain the same job title at both Firm 1 and Firm 2 (i.e., CEO in both firms, CFO in both firms, or General Counsel in both firms), we still find that adding the manager-specific fixed effects increases the models' adjusted R^2 by an average of 8.3 raw percentage points across the five disclosure attributes.

TABLE 3
Testing Individual Top Managers' Fixed Effects on Forecast Precision

F_Precision (n = 2,169)[a]	Economic Determinants, Year- and Firm-Specific Fixed Effects (Model 1)	Economic Determinants *and* Manager Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 2)	Economic Determinants *and* CEOs' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 3)	Economic Determinants *and* CFOs' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 4)	Economic Determinants *and* General Counsels' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 5)
Testing Economic Determinants = 0					
F-statistics	1.33	1.37	1.31	1.40	1.23
(p-value)	(0.172)	(0.150)	(0.185)	(0.132)	(0.237)
Constraints	16	16	16	16	16
Testing Manager Fixed Effects = 0					
F-statistics		34.88	12.20	21.12	8.69
(p-value)		(<0.001)	(<0.001)	(<0.001)	(<0.001)
Constraints		235	138	84	17
Adjusted R^2	1.8%	10.0%	8.3%	3.8%	1.7%
Improvement in Adjusted R^2 Relative to Model (1):					
In Raw Percentage Points		8.2	6.5	2.0	−0.1
As a % of Model 1's R^2		456%	361%	111%	−6%

Distribution of Manager Fixed Effects[b] (n = 235 Managers)

	Manager Fixed Effect
% Significant at 10% Level	39%
Mean Effect	0.04
Median Effect	0.02
25th Percentile	−0.39
75th Percentile	0.36
Difference between 75th and 25th Percentile	0.75

[a] Exhibit 1 defines the economic determinants. Footnotes in Table 2 explain the statistical tests.

[b] The manager fixed effects are the estimated coefficients from Model 2 that controls for economic determinants as well as year- and firm-specific fixed effects in estimating the *incremental* manager fixed effects.

Table 4, replacing a manager at the 25th percentile with one at the 75th shifts away from a tendency to issue forecasts to convey bad news, toward a tendency to issue forecasts to convey good news. Table 5 shows that replacing a manager in the 25th percentile with one in the 75th shifts from forecasts that undershoot earnings to forecasts that overshoot earnings. Finally, Table 6 shows that replacing a manager in the 25th percentile with one in the 75th increases average absolute forecast error by 0.91 percent of stock price, which is material relative to the sample mean absolute error of 0.58 percent per Table 1.

Taken together, Tables 2–6 show that manager-specific fixed effects play an economically significant incremental role in explaining forecast frequency as well as detailed forecast attributes such as precision, news conveyed by the forecast, bias relative to *ex post* actual EPS, and absolute

TABLE 6

Testing Individual Top Managers' Fixed Effects on Absolute Forecast Error

F_AbsError (n = 1,914)[a]	Economic Determinants, Year- and Firm-Specific Fixed Effects (Model 1)	Economic Determinants *and* Manager Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 2)	Economic Determinants *and* CEOs' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 3)	Economic Determinants *and* CFOs' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 4)	Economic Determinants *and* General Counsels' Fixed Effects, Year- and Firm-Specific Fixed Effects (Model 5)
Testing Economic Determinants = 0:					
F-statistics	6.72	5.82	6.37	5.95	6.84
(p-value)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Constraints	16	16	16	16	16
Testing Manager Fixed Effects = 0:					
F-statistics		51.50	6.75	13.78	3.12
(p-value)		(<0.001)	(<0.001)	(<0.001)	(<0.001)
Constraints		213	126	76	15
Adjusted-R^2	12.5%	22.8%	19.4%	14.0%	14.0%
Improvement in Adjusted R^2 Relative to Model (1):					
In Raw Percentage Points		10.3	6.9	1.5	1.5
As a % of Model 1's R^2		82%	55%	12%	12%

Distribution of Manager Fixed Effects[b] (n = 213 Managers)

	Manager Fixed Effect
% Significant at 10% Level	44%
Mean Effect	0.03
Median Effect	−0.00
25th Percentile	−0.43
75th Percentile	0.48
Difference between 75th and 25th Percentile	0.91

[a] Exhibit 1 defines the economic determinants. Footnotes in Table 2 explain the statistical tests.

[b] The manager fixed effects are the estimated coefficients from Model 2 that controls for economic determinants as well as year- and firm-specific fixed effects in estimating the *incremental* manager fixed effects.

deviations. That is, do managers take the initiative to impose their styles on their firms, or do firms hire managers because the firm wants the manager to bring his/her unique disclosure style (which we refer to as the matching story)? Under either scenario, managers influence the firm's disclosure choices. The only difference is the extent of the firm's role in proactively hiring a manager with a particular disclosure style.

We argue that the impetus likely stems from top managers independently imposing their styles on the firm. First, Bertrand and Schoar (2003) conclude that the matching story does not explain the significant manager-specific fixed effects on fundamental strategic and operational choices such as preferences for internal versus external growth (e.g., capital investments and R&D versus acquisitions) and financial aggressiveness (e.g., dividend and capital structure policies). If firms do

not proactively hire managers for specific styles in fundamental investment and operational decisions, it stretches credulity that firms proactively hire new top executives—particularly CEOs—largely to change attributes of voluntary management earnings forecasts. Second, respondents to a recent National Investor Relations Institute (NIRI) survey say top managers do impose their unique disclosure styles on the firm.[21] Third, the matching story would require the firm's desired disclosure strategy to change over time in a pattern directly mirrored by changes in the types of managers the firm hires, and the primary causes of changes in desired disclosure strategy would have to be new drivers of voluntary disclosure that: (1) are not firm-specific characteristics (which are controlled by the firm-specific fixed effects), (2) are not year-specific effects (which are controlled by the year fixed effects), and (3) have not been identified by prior research, and thus are not included in our vector of economic determinants.

Testing Whether Top Managers' Personal Demographic Characteristics are Associated with Their Unique Disclosure Styles

Table 7 provides descriptive statistics on our sample managers' demographic characteristics. Nearly half have an accounting or finance background, 8 percent have a legal background, and the remainder rose from general management (including production and marketing). Of the other demographic characteristics we measure, 16 percent of our sample managers were born before World War II, 10 percent have military experience, and 37 percent hold an M.B.A.. We also manually collected demographic characteristics of a random sample of 100 managers from the Execucomp/First Call population of firms used for comparison in Table 1. Our sample managers are not significantly different from the random sample, except that we have a smaller proportion of older managers, which would bias against finding results on our age cohort variable (*PreWWII*). Panel B shows that managers rising from accounting and finance are less likely to be born before World War II, and more likely to hold M.B.A.s, consistent with a trend toward choosing professional administrators who run the company "by the numbers."

Table 8 presents the results of estimating Model 6 that examines the association between managers' disclosure styles and their demographic characteristics. The base group impounded in the intercept is managers born after World War II who have neither military experience nor an M.B.A., and who rose from general management. Table 8 shows that managers with accounting or finance backgrounds develop disclosure styles that reflect lower tolerance of ambiguity (fewer but more specific forecasts, $p \leq 0.05$) and that are conservative in tending to undershoot actual earnings ($p = 0.04$). Managers with legal backgrounds favor disclosure styles that guide expectations down ($p = 0.09$), reflecting sensitivity to litigation risk. Those born before World War II are conservative in that they tend to issue fewer forecasts ($p < 0.01$). Managers with military experience tend to develop disclosure styles that: (1) are more precise ($p = 0.04$), reflecting lower tolerance of ambiguity, and (2) provide downward (not upward) guidance ($p = 0.02$). Even though M.B.A.s' disclosure styles tend to guide expectations up ($p = 0.05$), their disclosures are more accurate relative to *ex post* actual EPS ($p = 0.08$).

These associations between managers' demographic characteristics and their fixed effects are of interest in their own right. Finding plausible associations between manager-specific fixed effects estimated in the first stage and distinctive (permanent) characteristics of their own personal back-

[21] Among firms that discontinued guidance, the most popular response to "What would make you consider providing earnings guidance in the future?" was "change in management philosophy." Similarly, when asked "Where did the decision to consider discontinuing guidance come from?" 69 percent responded that it came from senior management. For companies that are currently providing guidance, but are considering discontinuing guidance, 63 percent of the respondents indicated that this is attributable to a "change in management philosophy" (NIRI 2007).

TABLE 7

Descriptive Statistics on Managers' Personal Demographic Characteristics[a]

Panel A: Frequency of Managers' Demographic Characteristics

Demographic Characteristics (n = 292)	Frequency	Percent	Execucomp/First Call Random Sample Percent
Functional Career Track			
Acct/Fin	135	46	38
Legal	22	8	12
General Management	135	46	50
Age Cohort			
PreWWII	48	16	31*
Boomer/Xer	244	84	69*
Military Experience			
Military	29	10	6
Non-military	263	90	94
M.B.A. Education			
M.B.A.	107	37	29
Non-M.B.A.	185	63	71

Panel B: Age Cohort, Military Experience, and M.B.A. Degree by Functional Career Track

	Functional Career Track			
	Acct/Fin	Legal	General Management	Total
By Age Cohort				
PreWWII	15	4	29	48
Boomer/Xer	120	18	106	244
Total	135	22	135	292
Chi-Square Test of Independence: χ^2 (2) = 5.34, p = 0.07				
By Military Experience				
Military	11	3	15	29
Non-military	124	19	120	263
Total	135	22	135	292
Chi-Square Test of Independence: χ^2 (2) = 1.03, p = 0.60				
By M.B.A. Status:				
M.B.A.	68	0	39	107
Non-M.B.A.	67	22	96	185
Total	135	22	135	292
Chi-Square Test of Independence: χ^2 (2) = 27.18, p < 0.001				

* Indicates our sample differs significantly from the random sample managers (n = 100) selected from the Execucomp/First Call population on a two-sample test of differences in proportion ($p < 0.05$).

[a] We collect managers' functional career track (i.e., employment history), age, military experience, and education from Marquis Who's Who Online, Hoover's, Mergent Online, Zoominfo.com, NNBD.com, Execucomp, and general Internet searches. We classify a manager's functional background into one of three categories: Accounting and Finance (Acct/Fin), Legal, or General Management (including production, marketing, and strategy).

TABLE 8

Testing Associations between Demographic Characteristics of Managers' Personal Backgrounds and Their Own Unique Overall Disclosure Styles[a]

Model: $\lambda_{m,Yi} = \alpha_0 + \gamma_1 Acct/Fin + \gamma_2 Legal + \gamma_3 PreWWII + \gamma_4 Military + \gamma_5 MBA + \epsilon$

	F_Freq	*F_Precision*	*F_Goodnews*	*F_BiasUP*	*F_AbsError*
Acct/Fin = 1	−0.33	0.22	0.02	−0.22	0.16
	(0.05)**	(0.01)**	(0.78)	(0.04)**	(0.15)
Legal = 1	−0.16	0.18	−0.17	−0.15	0.17
	(0.65)	(0.16)	(0.09)*	(0.43)	(0.50)
PreWWII = 1	−0.70	−0.06	−0.10	−0.03	0.15
	(0.00)***	(0.60)	(0.30)	(0.83)	(0.40)
Military = 1	0.11	0.29	−0.22	0.05	−0.00
	(0.67)	(0.04)**	(0.02)**	(0.72)	(0.99)
MBA = 1	0.06	0.01	0.13	−0.13	−0.19
	(0.72)	(0.91)	(0.05)*	(0.20)	(0.08)*
Adjusted R^2 (OLS)	3%	5%	7%	2%	1%
Robust Adjusted R^2 (LTS)[b]	12%	13%	5%	9%	11%

*,**,*** Significant at 10 percent, 5 percent, and 1 percent, respectively.

Two-tailed robust p-values adjusting for heteroscedasticity appear in parentheses.

[a] $\lambda_{m,Yi}$ are the manager-specific fixed effects estimated in Model 2. *Acct/Fin*, *Legal*, *PreWWII*, *Military*, and *MBA* are dichotomous variables characterizing whether managers rise from accounting/finance or legal backgrounds, whether they were born before World War II, whether they have military experience, and whether they hold an M.B.A. degree. To ensure our inferences are not an artifact of a few extreme values, we report OLS results after removing a small number of extreme outliers with (absolute) studentized residuals exceeding 2.

[b] Robust adjusted R^2 is generated from the least trimmed squares robust regression (Rousseeuw 1984).

grounds also confirms these fixed effects are not simply random, but instead capture systematic long-lived cross-sectional differences in managers' unique disclosure styles.

The explanatory power of these demographic characteristics is moderate, however.[22] Two econometric issues dampen the adjusted R^2 from OLS estimation. First, the dependent variables in Model 6 (i.e., manager-specific disclosure styles) are estimated parameters from the fixed-effect analysis, and so contain measurement error (although this measurement error should be smaller than the error that would occur in simpler designs that do not cleanly separate the manager effect from the firm effect). While OLS estimation still generates unbiased coefficients (Wooldridge

[22] While moderate, our adjusted R^2s compare favorably to those of the only other study (of which we are aware) that reports statistics on the explanatory power of models attempting to explain cross-sectional variation in managers' fixed effects (or styles) regarding other corporate decisions. In an independent and contemporaneously developed study, Dyreng et al. (2010, 1185) document manager-specific effects on *corporate tax avoidance* (measured based on effective tax rates). They conclude that "the results indicate that biographical information such as educational background and age does not explain much of the variation in tax avoidance across executives" because adjusted R^2s from their regressions are negative. However, the authors go on to point out that "this finding does not suggest that executives have no effect on tax avoidance. Instead, these results are evidence that common, observable characteristics are not strongly associated with executives' propensities to reduce effective tax rates."

2006, 320), this larger error variance decreases R^2. We ran a simulation to confirm that measurement error in the dependent variable can significantly dampen adjusted R^2.[23]

Second, measurement error in the dependent variable can lead to outliers. We repeat our analysis using robust regression via least trimmed squares, which generates more stable results in the presence of outliers (Chen 2002; Zaman et al. 2001). The last row of Table 8 shows that for four of our five forecast characteristics, the demographic variables' explanatory power is quite a bit higher using the robust regression estimation, with adjusted R^2s ranging from 5 percent to 13 percent.

The simulation and robust regression results give a more appropriate sense of the economic significance of the demographic characteristics—the recommended starting point for exploring differences in managers' unique styles (Finkelstein and Hambrick 1996; Hambrick 2007, 335). Although we conducted a comprehensive review of a broad range of literatures, theory as to how demographic characteristics might be associated with unique managerial styles is in its infancy, and there are a limited number of observable (permanent) demographic characteristics that could plausibly affect disclosure. Managers' individual-specific disclosure styles undoubtedly result from their many (unobservable) experiences, psychological traits, and values, for which demographic characteristics such as functional career track, age cohort, military experience, and M.B.A. education provide a start, but are admittedly an incomplete and crude set of proxies.

Additional Analyses

Given the modest explanatory power of the demographic variables, we also explore whether two nondemographic factors explain managers' disclosure styles: managers' risk-taking behavior as revealed by their fixed effects on their firms' operating and financing decisions, and managers' fixed effects on their firms' operational efficiency.[24] We identify firm decisions and outcomes that might reflect their managers' risk-taking behavior: the frequency of acquisitions, investment in R&D and capital expenditure, cash holdings, and stock return volatility (Bertrand and Schoar 2003; Bargeron et al. 2010). We estimate managers' fixed effects on these decisions and outcomes using models similar to Bertrand and Schoar (2003). We add these risk-related manager-specific fixed effects (individually and in a composite index) to our original set of demographic variables in Model 6 that explains managers' overall disclosure styles. Our original inferences remain largely robust, and we find little evidence that these risk-related manager-specific fixed effects are associated with managers' styles in voluntary disclosure.[25]

---

[23] Even in a simulated world where X fully explains the true value of Y (so R^2 is 100 percent absent measurement error), adjusted R^2 plummets from 97 percent when the standard deviation of the error term in Y is 10 percent of the mean of Y, to just 11 percent when the standard deviation of the error term increases to 1.5 times the mean of Y. (Specifically, we randomly generate 200 observations by drawing X from a uniform distribution [1,100]. Y is determined by X, $Y = 1+1 * X$, so R^2 is 100 percent when we regress Y on X. We also randomly draw 200 values of a measurement error term, v, from a normal distribution with mean zero and standard deviation that is 10 percent of the mean of Y. We construct $Y^* = Y + v$ and regress Y^* on X. We repeat this regression 1,000 times (each time randomly drawing 200 values for the error term v from the same distribution), constructing Y^* based on v each time, to get the mean adjusted R^2 and the mean coefficient estimates on X. We repeat this simulation for 15 trials, and in each trial we increase the standard deviation of the error term v by 10 percent of the mean of Y until the standard deviation is 1.5 times the mean of Y.)

[24] We are able to identify one demographic characteristic that likely primarily reflects risk aversion. Research in psychology and economics suggests that women tend to be more risk-averse, both in general (Byrnes et al. 1999; Harris et al. 2006) and in financial decisions (Powell and Ansic 1997; Watson and McNaughton 2007; Sapienza et al. 2009). When we include gender in Model 6, all of the inferences on our original demographic variables remain robust, and we find that female managers tend to adopt disclosure styles that are less precise and underestimate upcoming earnings, consistent with females' greater risk aversion. However, recall that only six of our managers are female, so our findings on gender are necessarily tentative.

[25] The only exception is that managers who are more risk-seeking in terms of engaging in more frequent acquisitions tend to forecast more frequently and tend to be more precise. When we include this fixed effect in our Model 6, the

There are several explanations why this additional analysis did not yield new insights. First, to the extent our original demographic variables (partially) reflect the risk preferences of demographic groups (e.g., those who are older and those rising from accounting backgrounds tend to be more risk-averse), the incremental effect picked up by the new risk variables does not reflect the manager's total risk preference, but rather reflects how the manager's risk preference deviates from the average risk preferences of his/her demographic classifications. Second, unlike the reliably observable demographic variables, the new measures are *estimated*. Third, estimates of manager-specific fixed effects on the firm's operational and financing decisions may not reflect managers' true risk preferences because firms often provide compensation (e.g., equity-based pay) to encourage even risk-averse managers to take on risky operational and financing projects to maximize firm value (Rajgopal and Shevlin 2002; Coles et al. 2006).

With respect to managers' operational efficiency, we estimate managers' fixed effects on asset turnover using a model similar to Bertrand and Schoar (2003). We find only weak evidence that managers' operational efficiency is associated with one aspect of their disclosure styles—managers who are associated with higher asset turnover tend to be less precise ($p = 0.09$), while the inferences on our original demographic variables remain robust.

Recall that Tables 2–6 document the incremental effect of individual managers' *total* influences on cross-sectional variation in firms' voluntary disclosure choices in order to address our first research question: Do individual managers exert economically significant idiosyncratic non-economic influence on their firms' voluntary disclosure choices? Our second research question follows upper echelons theory's recommendation to focus on how managers' demographic characteristics affect their unique styles, by investigating whether reliable, observable demographic characteristics of managers' personal backgrounds help explain cross-sectional variation in *managers'* own unique voluntary disclosure styles. Table 8 provides evidence of how managers' demographic characteristics affect their overall disclosure styles identified in Tables 2–6.

We now provide ancillary analysis to address a different question—do the necessarily limited set of demographic characteristics for which we can find some conceptual support in the literature play an incremental role in directly explaining cross-sectional variation in *firms'* voluntary disclosure choices? We replace the manager-specific indicator variables used in Model 2 to estimate managers' *total* disclosure styles with the set of their demographic characteristics. Regressing firm-level disclosure choices directly on this admittedly limited set of demographic characteristics captures only the portion of the manager's total disclosure style that is associated with these demographic characteristics. Nonetheless, Table 9 shows that seven of the nine significant results reported in Table 8 continue to hold (accounting/finance backgrounds are no longer associated with less frequent but more downward-biased forecasts).[26]

---

coefficient on *Legal* in the analysis of *F_Goodnews* just loses significance ($p = 0.14$).

[26] Table 9 does not provide evidence on managers' *total* impact on their firms' disclosure decisions; instead it only captures the effect of a set of observable demographic characteristics that is necessarily limited by the dearth of both extant theory and data availability. Thus, Table 9 captures only a small portion of managers' total idiosyncratic effects on cross-sectional variation in firms' voluntary disclosure choices. In contrast, the primary specifications in Tables 2–6 document the *full* effect of individual managers on firms' disclosure choices, and then Table 8 zooms in to explore how their demographic characteristics influence their own unique overall disclosure styles.

TABLE 9

Testing Direct Associations Between Demographic Characteristics of Managers' Personal Backgrounds and Firms' Earnings Forecast Characteristics[a]

	F_Freq	*F_Precision*	*F_Goodnews*	*F_BiasUP*	*F_AbsError*
Acct/Fin = 1	0.02	0.12	−0.01	−0.06	0.07
	(0.84)	(0.09)*	(0.87)	(0.50)	(0.39)
Legal = 1	0.35	0.09	−0.22	0.10	0.33
	(0.14)	(0.52)	(0.02)**	(0.51)	(0.21)
PreWWII = 1	−0.33	−0.20	0.11	0.17	0.05
	(0.03)**	(0.08)*	(0.21)	(0.21)	(0.84)
Military = 1	−0.09	0.21	−0.29	0.03	0.04
	(0.67)	(0.08)*	(0.00)***	(0.83)	(0.85)
MBA = 1	0.20	−0.09	0.10	0.03	−0.15
	(0.07)*	(0.27)	(0.07)*	(0.77)	(0.08)*
Full Model Adjusted R^2	33.7%	2.1%	12.5%	12.4%	12.6%
Incremental Adjusted R^2 (relative to Model 1)	0.2%	0.3%	0.9%	−0.1%	0.1%

*, **, *** Significant at 10 percent, 5 percent, and 1 percent, respectively.

Two-tailed robust p-values adjusting for heteroscedasticity appear in parentheses.

[a] All variables are defined in Exhibit 1. This analysis replaces the manager-specific indicator variables that we used in Model 2 to estimate managers' *total* disclosure styles (as reported in Tables 2 through 6) with the set of their demographic characteristics. Consequently, this analysis captures only the portion of the manager's total disclosure style that is associated with these demographic characteristics. For brevity, we do not report coefficient estimates for the economic determinants, firm fixed effects, and time-period fixed effects.

V. CONCLUSIONS

Driven by conflicting views on whether managers impose any idiosyncratic influence on corporate decisions (e.g., Weintraub 2002; Feltham 2005; DiMaggio and Powell 1983; Hambrick and Mason 1984), our study first examines whether top managers exert an economically significant influence on their firms' voluntary disclosure choices. We estimate the incremental effect an individual manager exerts on voluntary disclosure over two or more employing firms, after controlling for firm fixed effects, time-specific fixed effects, and economic determinants of voluntary disclosure identified in prior research. From the increase in adjusted R^2, the proportion of significant manager fixed effects, and the interquartile range of the manager fixed effects, we conclude that managers exert economically significant individual-specific influence over five attributes of management earnings forecasts: the frequency of forecasts, forecast precision, news conveyed by the forecast, and the bias in and accuracy of the forecasts.

We then build on upper echelons theory (Hambrick and Mason 1984) and a broad set of literatures ranging from strategic management, to career counseling, to sociology, to psychology, to military science, and to business education to guide our exploration of how managers' functional career track, age cohort, military experience, and M.B.A. degree influence their unique disclosure styles. We find that disclosure styles of managers promoted from accounting and finance reflect elements of conservatism and distaste for ambiguity (fewer but more precise disclosures that tend to underestimate upcoming actual earnings). Disclosure styles of managers from legal backgrounds favor downward guidance, consistent with sensitivity to litigation risk. Managers born before World War II develop more conservative disclosure styles in terms of being more reluctant to forecast. Those with military experience favor more precise disclosure styles that tend

toward prompt disclosure of unfavorable information. Finally, managers with M.B.A.s develop styles that tend to guide expectations upward, but are more accurate, consistent with M.B.A.s being better at forecasting or more sensitive to the adverse consequences of inaccurate forecasts.

These plausible associations between the manager-specific fixed effects and distinctive permanent demographic characteristics of managers' own personal backgrounds confirm that the estimated manager-specific fixed effects reflect systematic long-lived differences in managers' disclosure styles, and not simply random error. From a practical perspective, evidence on the relation between managers' disclosure styles and their demographic characteristics can help users evaluate managers' disclosure choices. Knowing that older managers are less likely to issue forecasts, and that those with military experience are less likely to guide *unless* there is bad news, can help market participants interpret when silence might *not* indicate poor performance (contrary to theory that nondisclosure signals bad news [e.g., Verrecchia 2001, 141]).

Our evidence that the largely unexplored personal style dimension plays a significant incremental role in explaining cross-sectional differences in corporate voluntary disclosure choices responds to the Brochet et al. (2009, 32) suggestion for research on whether firm disclosure policies vary systematically with managerial styles. Our evidence also helps address the puzzle noted by Hirst et al. (2008, 316): "Given that managers have substantially greater control over forecast characteristics ... it is striking that the decisions managers make about such characteristics are comparatively less understood." Consistent with their conjecture that managers have more control over the detailed forecast characteristics, our evidence suggests that individual managers' idiosyncratic styles play a relatively more important incremental role in explaining detailed characteristics of the forecasts, such as precision. Our evidence also contributes in a broader fashion to the nascent manager-effects literature that has largely focused on manager-specific effects on high-profile strategic operating and financing decisions. Finding that managers also matter even in seemingly second-order decisions such as detailed voluntary disclosure choices extends the range of corporate decisions in which managers' idiosyncratic noneconomic styles are known to play an economically significant role. We also contribute by integrating a broad set of literatures to provide conceptual grounding as to why and when managers might (or might *not*) matter, and how various demographic characteristics of managers' personal backgrounds likely affect their idiosyncratic disclosure styles.

We believe future theoretical, archival, and experimental research incorporating richer roles for individuals' styles has great potential for increasing our understanding of how firms make financial reporting choices. For example, recent theory is beginning to allow idiosyncratic stylistic differences across individual top executives, such as managers' vision for the future of their companies, to affect corporate outcomes (e.g., Rotemberg and Saloner 2000; Van den Steen 2005). We hope that our evidence likewise spurs development of disclosure theories that incorporate richer roles for differences across managers' (noneconomic) characteristics.

Upper echelons theory suggests that managers' demographic characteristics are the appropriate starting point for exploring reasons for differences in individual managers' styles. Drawing on a comprehensive review of a broad set of literatures, we posit and find plausible associations between managers' demographic characteristics and their disclosure styles, although the explanatory power of the demographic characteristics is moderate. Consistent with upper echelon theory's emphasis on demographics, our attempts to identify nondemographic determinants of managers' styles met with limited success. Significant further improvements in explaining managers' styles may therefore require finer demographic data (relative to the types of demographic data that are currently publicly available), theoretically guided exploration of variables beyond demographic characteristics, and/or very different research methods. For example, surveys or behavioral research identifying cognitive traits and values that distinguish differences among the styles of highly successful professional managers may be fruitful avenues for future research.

APPENDIX
ECONOMIC DETERMINANTS OF VOLUNTARY DISCLOSURE
(the Vector X_{it})

To isolate the effects of individual managers, we control for the following comprehensive set of economic determinants that have been well documented in the prior disclosure literature. Prior research has established that firms are less likely to issue forecasts when earnings are more uncertain and harder to predict (e.g., Ajinkya and Gift 1984; Baginski et al. 2002; Ajinkya et al. 2005). Similar to that prior research, we control for the absolute value of the change in EPS deflated by lagged stock price ($|\Delta EPS|$) and the dispersion in analysts' forecasts scaled by the absolute value of the median forecast (*Disp*).[27] Firms missing either their prior period earnings or zero profit benchmarks are less likely to issue forecasts (e.g., Miller 2002), so we control for the sign of the change in the firm's EPS: *EPS_UP* equals 1 if EPS is greater than or equal to the prior year's EPS. We also include an indicator variable that equals 1 if the firm reports a loss (*Loss*).

Firms with higher proprietary information costs are less likely to issue public earnings forecasts, so we control for two measures of proprietary information costs: *R&D*, the firm's R&D expenditures scaled by total assets (Wang 2007), and *Mkt-Bk,* the firm's market-to-book ratio (Bamber and Cheon 1998; Ajinkya et al. 2005). Ajinkya et al. (2005) show that two measures of higher quality governance—proportion of outside directors and institutional ownership—are associated with a greater propensity to issue forecasts, so we control for the percentage of board members who are *not* also officers of the firm (*Outside_Dir*), and the percentage of the firm's common stock held by institutions (*%Inst*). To control for the fact that larger, more heavily followed firms issue more voluntary disclosures (e.g., Lang and Lundholm 1993; Ajinkya et al. 2005), we include the natural log of the market value of equity (*LSize*) and the number of analysts following the firm (*#Analysts*). Litigation risk is associated with voluntary disclosure. Following Wang (2007), we include an indicator variable that equals 1 if the firm is in a high litigation risk industry[28] *and* suffers an earnings decrease of 20 percent or more (*Litigation*).[29] Following Brochet et al. (2009), we control for lawsuits (*ClassAction* equals 1 if the firm is involved in a securities lawsuit in year *t*), restructurings (*Restructuring* equals 1 if the firm was engaged in a restructuring during year *t*), restatements (*Restatement* equals 1 if the firm had a restatement in year *t*), and acquisitions (*Acquisition* equals 1 if the firm has a merger or acquisition in year *t*).

As Hirst et al. (2008) note, much less is known about the determinants of detailed attributes of management forecasts. To be conservative, we include all of the above variables plus forecast horizon (e.g., Baginski and Hassell 1997; Bamber and Cheon 1998; Choi and Ziebart 2004) in the models explaining forecast precision, news conveyed by the forecast, and the bias and error in the

---

[27] Following Baginski et al. (2002), we control for the difficulty of predicting current earnings, since this is most relevant to the decision whether to issue an earnings forecast in the current period. These measures also avoid survivorship bias induced by historical measures of earnings predictability.

[28] Following Francis et al. (1994) and Wang (2007), we classify the following as high litigation risk industries: SIC codes 2833–2836 (biotechnology), 3570–3577 and 7370–7374 (computers), 3600–3674 (electronics), 5200–5961 (retailing), and 8731–8734 (R&D service).

[29] Our model controls for over-time changes in the regulatory environment through the year-specific fixed effects. There is limited evidence that two additional phenomena may be associated with the propensity to issue management forecasts (although there is no evidence these phenomena are related to detailed attributes of management forecasts). Nagar et al. (2003) conclude that firms whose managers enjoy larger stock-based incentives issue more earnings forecasts (although Barth [2003] questions whether the results might reflect a correlated omitted variable). There is also conflicting evidence whether insider trading is associated with the frequency of management earnings forecasts. Noe (1999) finds no evidence that managers profit by trading before public disclosures, and Cheng and Lo (2006) find managers do *not* adjust their forecasting activity when selling shares (by far the most common type of insider trading), but they find some evidence that managers increase the number of bad news forecasts prior to share purchases. Untabulated results reveal that adding controls for stock-based incentives (i.e., the proportion of managers' compensation that is sensitive to share price, and managerial shareholdings), and total insider buying does not affect our inferences.

forecast relative to *ex post* actual EPS. We define *F_Horizon* as the number of days between the forecast announcement and the end of the fiscal period, divided by 360 days for forecasts of annual earnings, and divided by 90 days for forecasts of quarterly earnings. Thus, *F_Horizon* measures the proportion of the fiscal period remaining at the date of the management forecast. Exhibit 1 summarizes our variable definitions.

REFERENCES

Ajinkya, B., S. Bhojraj, and P. Sengupta. 2005. The association between outside directors, institutional investors and the properties of management earnings forecasts. *Journal of Accounting Research* 43 (3): 343–376.

———, and M. Gift. 1984. Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research* 22 (2): 425–444.

Anilowski, C., M. Feng, and D. Skinner. 2007. Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. *Journal of Accounting and Economics* 44 (1–2): 36–63.

Baginski, S., and J. Hassell. 1997. Determinants of management forecast precision. *The Accounting Review* 72 (2): 303–312.

———, ———, and M. Kimbrough. 2002. The effect of legal environment on voluntary disclosure: Evidence from management earnings forecasts issued in U.S. and Canadian markets. *The Accounting Review* 77 (1): 25–50.

Bagley, C. 2008. Winning legally: The value of legal astuteness. *Academy of Management Review* 33 (2): 378–390.

Bamber, L., and Y. Cheon. 1998. Discretionary management earnings forecast disclosures: Antecedents and outcomes associated with forecast venue and forecast specificity. *Journal of Accounting Research* 36 (2): 167–190.

Bargeron, L., K. Lehn, and C. Zutter. 2010. Sarbanes-Oxley and corporate risk-taking. *Journal of Accounting and Economics* 49 (1–2): 34–52.

Barth, M. 2003. Discussion of: Compensation policy and discretionary disclosure. *Journal of Accounting and Economics* 34 (January): 311–318.

Bertrand, M., and A. Schoar. 2003. Managing with style: The effects of managers on firm policies. *The Quarterly Journal of Economics* 118 (4): 1169–1208.

Bhargava, S., and F. Jespersen. 1993. Portrait of a CEO. *BusinessWeek* (October 11): 64–65.

Blanchard, O. 1993. Movements in the equity premium. In *Brookings Papers on Economic Activity*, edited by Shiller, R., and J. Siegel, 75–138. Washington, D.C.: Brookings Institution Press.

Brochet, F., L. Faurel, and S. McVay. 2009. Earnings guidance following top executive turnovers. Working paper, Harvard Business School.

Byrnes, J., D. Miller, and W. Schafer. 1999. Gender differences in risk taking: A meta-analysis. *Psychological Bulletin* 125 (3): 367–383.

Cameron, A., and P. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York, NY: Cambridge University Press.

Chalmers, K., and J. Godfrey. 2004. Reputation costs: The impetus for voluntary derivative financial instrument reporting. *Accounting, Organizations and Society* 29 (2): 95–125.

Chen, C. 2002. Robust regression and outlier detection with the ROBUSTREG procedure. *Proceedings of the Twenty-Seventh Annual SAS Users Group International Conference*, Cary, NC, SAS Institute Inc.

Chen, S. 2004. Why do managers fail to meet their own forecasts? Working paper, University of Texas.

Cheng, Q., and K. Lo. 2006. Insider trading and voluntary disclosures. *Journal of Accounting Research* 44 (5): 815–848.

Choi, J., and D. Ziebart. 2004. Management earnings forecasts and the market's reaction to predicted bias in the forecast. *Asia-Pacific Journal of Accounting and Economics* 11 (2): 167–192.

Christensen, P., and G. Feltham. 2003. *Information in Markets*. Economics of Accounting, Vol. I. Boston, MA: Kluwer.

Coles, J., D. Naveen, and L. Naveen. 2006. Managerial incentives and risk-taking. *Journal of Financial Economics* 79 (2): 431–468.

Crossland, C., and D. Hambrick. 2007. How national systems differ in their constraints on corporate executives: A study of CEO effects in three countries. *Strategic Management Journal* 28 (8): 767–789.

Datz, T. 2000. Equity? *CIO* (January 15): 74–84.

DiMaggio, P., and W. Powell. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* 48 (2): 147–160.

Dye, R. 1986. Proprietary and nonproprietary disclosures. *The Journal of Business* 59 (2): 331–366.

Dyreng, S., M. Hanlon, and E. Maydew. 2010. The effects of executives on corporate tax avoidance. *The Accounting Review* 85 (4): 1163–1189.

Evans, J., and S. Sridhar. 2002. Disclosure-disciplining mechanisms: Capital markets, product markets, and shareholder litigation. *The Accounting Review* 77 (3): 595–626.

Faircloth, A., and D. Ricchiute. 1981. Ambiguity intolerance and financial reporting alternatives. *Accounting, Organizations and Society* 6 (1): 53–67.

Feltham, G. 2005. Information economics and management accounting: A brief personal perspective. *Journal of Management Accounting Research* 17: 145–152.

Finkelstein, S., and D. Hambrick. 1996. *Strategic Leadership: Top Executives and Their Effects on Organizations*. St. Paul, MN: West Publishing Company.

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32 (2): 137–164.

Franke, V. 1998. Old ammo in new weapons? Comparing value-orientation of experienced and future military leaders. *Journal of Political and Military Sociology* 26 (2): 253–271.

———. 2001. Generation X and the military: A comparison of attitudes and values between West Point cadets and college students. *Journal of Political and Military Sociology* 29 (1): 92–119.

Ge, W., D. Matsumoto, and J. Zhang. 2009. Do CFOs have styles of their own? An empirical investigation of the effect of individual CFOs on financial reporting practices. Working paper, University of Washington.

Ghoshal, S. 2005. Bad management theories are destroying good management practices. *Academy of Management Learning & Education* 4 (1): 75–91.

Gintis, H., and R. Khurana. 2008. Corporate honesty and business education: A behavioral model. In *Moral Markets: The Critical Role of Values in the Economy*, edited by Zak, P., 300–327. Princeton, NJ: Princeton University Press.

Goertzel, T., and A. Hengst. 1971. The military socialization of university students. *Social Problems* 19 (2): 258–267.

Hambrick, D. 2007. Upper echelons theory: An update. *Academy of Management Review* 32 (2): 334–343.

———, M. Geletkanycz, and J. Fredrickson. 1993. Top executive commitment to the status quo: Some tests of its determinants. *Strategic Management Journal* 14 (6): 401–418.

———, and P. Mason. 1984. Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review* 9 (2): 193–206.

Hannan, M., and J. Freeman. 1977. The population ecology of organizations. *American Journal of Sociology* 82 (5): 929–964.

Harris, C., M. Jenkins, and D. Glaser. 2006. Gender differences in risk assessment: Why do women take fewer risks than men? *Judgment and Decision Making* 1 (1): 48–63.

Headquarters, Department of the Army. 1993. *FM 100–5 Operations*. Washington, D.C.: Government Printing Office.

Healy, P., A. Hutton, and K. Palepu. 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research* 16 (3): 485–520.

Hirst, E., L. Koonce, and S. Venkataraman. 2008. Management earnings forecasts: A review and framework. *Accounting Horizons* 22 (3): 315–338.

Hitt, M., and B. Tyler. 1991. Strategic decision models: Integrating different perspectives. *Strategic Management Journal* 12 (5): 327–351.

Holland, J. 1966. A psychological classification scheme for vocations and major fields. *Journal of Counseling Psychology* 13 (3): 278–288.

———. 1973. *Making Vocational Choices: A Theory of Careers*. Englewood Cliffs, NJ: Prentice Hall.

———. 1997. *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments*. Odessa, FL: Psychological Assessment Resources, Inc.

Jensen, M., and E. Zajac. 2004. Corporate elites and corporate strategy: How demographic preferences and structural position shape the scope of the firm. *Strategic Management Journal* 25 (6): 507–524.

Kilduff, M. 2006. Editor's comments: Prize-winning articles for 2005 and the first two decades of *AMR*. *Academy of Management Review* 31 (4): 792–793.

Lang, M. 1999. Discussion of: Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research* 16 (3): 521–524.

———, and R. Lundholm. 1993. Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research* 31 (2): 246–271.

———, and ———. 2000. Voluntary disclosure and equity offerings: Reducing information asymmetry or hyping the stock? *Contemporary Accounting Research* 17 (4): 623–662.

Lieberson, S., and J. O'Connor. 1972. Leadership and organizational performance: A study of large corporations. *American Sociological Review* 37 (2): 117–130.

Malmendier, U., and S. Nagel. 2008. Depression babies: Do macroeconomic experiences affect risk-taking? Working paper, University of California, Berkeley.

McCabe, D., K. Butterfield, and L. Trevino. 2006. Academic dishonesty in graduate business programs: Prevalence, causes, and proposed action. *Academy of Management Learning & Education* 5 (3): 294–305.

Miller, G. 2002. Earnings performance and discretionary disclosure. *Journal of Accounting Research* 40 (1): 173–204.

Nagar, V., D. Nanda, and P. Wysocki. 2003. Discretionary disclosure and stock-based incentives. *Journal of Accounting and Economics* 34 (1–3): 283–309.

Naranjo-Gil, D., and F. Hartmann. 2006. Management accounting systems, top management team heterogeneity and strategic change. *Accounting, Organizations and Society* 32 (7–8): 735–756.

National Investor Relations Institute (NIRI). 2007. *NIRI 2007 Earnings Guidance Practices Survey Results*. Vienna, VA: NIRI.

Noe, C. 1999. Voluntary disclosures and insider transactions. *Journal of Accounting and Economics* 27 (3): 305–326.

Nonis, S., and C. Swift. 2001. An examination of the relationship between academic dishonesty and workplace dishonesty: A multicampus investigation. *Journal of Education for Business* 77 (2): 69–77.

Palsson, A. 1996. Does the degree of relative risk aversion vary with household characteristics? *Journal of Economic Psychology* 17 (6): 771–787.

Pennebaker, J., and L. King. 1999. Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology* 77 (6): 1296–1312.

Powell, M., and D. Ansic. 1997. Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology* 18 (6): 605–628.

Rajgopal, S., and T. Shevlin. 2002. Empirical evidence on the relation between stock option compensation and risk taking. *Journal of Accounting and Economics* 33 (2): 145–171.

Rotemberg, J., and G. Saloner. 2000. Visionaries, managers, and strategic direction. *The Rand Journal of Economics* 31 (4): 693–716.

Rousseeuw, P. 1984. Least median of squares regression. *Journal of the American Statistical Association* 79 (388): 871–880.

Salanick, G., and J. Pfeffer. 1977. Constraints on administrator discretion: The limited influence of mayors on city budgets. *Urban Affairs Quarterly* 12 (4): 475–498.

Sapienza, P., L. Zingales, and D. Maestripieri. 2009. Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of National Academy of Sciences USA* 106 (36): 15268–15273.

Schuman, H., and J. Scott. 1989. Generations and collective memories. *American Sociological Review* 54 (3): 359–381.

Smith, M., and M. White. 1987. Strategy, CEO specialization, and succession. *Administrative Science Quarterly* 32 (2): 263–280.

Soeters, J. 1997. Value orientations in military academies: A thirteen country study. *Armed Forces and Society* 24 (1): 7–32.

Sorter, G., S. Becker, T. Archibald, and W. Beaver. 1964. Corporate personality as reflected in accounting decisions: Some preliminary findings. *Journal of Accounting Research* 2 (2): 183–196.

Spender, J. 1989. *Industry Recipes: An Enquiry into the Nature and Sources of Managerial Judgment.* Oxford, U.K.: Blackwell.

Thomas, A., R. Litschert, and K. Ramaswamy. 1991. The performance impact of strategy-manager coalignment: An empirical examination. *Strategic Management Journal* 12 (7): 509–522.

Tihanyi, L., A. Ellstrand, C. Daily, and D. Dalton. 2000. Composition of the top management team and firm international diversification. *Journal of Management* 26 (6): 1157–1177.

Van den Steen, E. 2005. Organizational beliefs and managerial vision. *Journal of Law Economics and Organization* 21 (1): 256–283.

Verrecchia, R. 2001. Essays on disclosure. *Journal of Accounting and Economics* 32 (1–3): 97–180.

Wang, I. 2007. Private earnings guidance and its implications for disclosure regulation. *The Accounting Review* 82 (5): 1299–1332.

Watson, J., and M. McNaughton. 2007. Differences in risk aversion and expected retirement benefits. *Financial Analysts Journal* 63 (4): 52–62.

Weintraub, E. 2002. Neoclassical Economics. In *The Concise Encyclopedia of Economics*, edited by D. Henderson. Indiannapolis, IN: Liberty Fund, Inc. Available from http://www.econlib.org/Library/Encl/NeoclassicalEconomics.html.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: The MIT Press.

——— 2006. *Introductory Econometrics: A Modern Approach.* 3rd edition. Mason, OH: Thomson South-Western.

Zaman, A., P. Rousseeuw, and M. Orhan. 2001. Econometric applications of high-breakdown robust regression techniques. *Economics Letters* 71 (1): 1–8.

Zemke, R., C. Raines, and P. Filipczak. 2000. *Generations at Work: Managing the Clash of Veterans, Boomers, Xers, and Nexters in Your Workplace.* New York, NY: Amacom.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1163–1189

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1163

The Effects of Executives on Corporate Tax Avoidance

Scott D. Dyreng
Duke University

Michelle Hanlon
Massachusetts Institute of Technology

Edward L. Maydew
University of North Carolina

ABSTRACT: This study investigates whether individual top executives have incremental effects on their firms' tax avoidance that cannot be explained by characteristics of the firm. To identify executive effects on firms' effective tax rates, we construct a data set that tracks the movement of 908 executives across firms over time. Results indicate that individual executives play a significant role in determining the level of tax avoidance that firms undertake. The economic magnitude of the executive effects on tax avoidance is large. Moving between the top and bottom quartiles of executives results in approximately an 11 percent swing in GAAP effective tax rates; thus, executive effects appear to be an important determinant in firms' tax avoidance.

Keywords: *tax avoidance; effective tax rate; executive fixed effects.*

Data Availability: *Data used in this study are available from public sources identified in the document.*

I. INTRODUCTION

This study investigates whether individual executives have incremental effects on their firms' tax avoidance that cannot be explained by characteristics of the firm. Despite decades of empirical research in corporate taxation, little attention has been focused on whether individual executives have an effect on their firms' tax avoidance (for reviews see Shackelford and Shevlin [2001] and Graham [2003]). In this prior literature, executives were either ignored or treated as homogenous inputs to the tax avoidance process. In contrast, our study

We thank Vic Anand, Mark Nelson, Tom Omer (editor), two anonymous referees, and workshop participants at The University of Arizona, Arizona State University, University of California, Berkeley, Cornell University, The University of Georgia, University of Notre Dame, University of Pennsylvania, and Stanford University. We thank Nemit Shroff, Jake Thornock, and the Kresge Library staff at the University of Michigan for assisting with the collection of the executive biographical data. Professor Dyreng acknowledges financial support from the Deloitte Doctoral Fellowship. Professor Maydew acknowledges financial assistance from the David E. Hoffman Chair at the University of North Carolina.

Editor's note: Accepted by Thomas Omer

Submitted: June 2008
Accepted: December 2009
Published Online: June 2010

considers the possibility that individual top executives are partially responsible for variation in tax avoidance across firms. Essentially, we posit that a given firm's tax avoidance will differ depending on whether it is run by Executive A versus Executive B. We find that not only do executives matter incremental to firm characteristics, but also they appear to matter in a big way.

At first thought, it might be hard to imagine a top executive having an individual effect on the firm's tax avoidance. The typical CEO is almost never a tax expert. While the typical CEO is unlikely to understand the ins and outs of common tax strategies, he/she likely understands the competitive nature of his/her industry and the potential for expansion to generate operational economies of scale. Thus, it is reasonable that a CEO could affect the firm's operational and financial strategies, but perhaps less so the firm's tax avoidance activities.

However, a CEO can affect tax avoidance by setting the "tone at the top" with regard to the firm's tax activities. For example, some CEOs may change the relative emphasis of different functional areas of the firm (e.g., marketing, operations, treasury, tax) and the resources allocated to hiring different advisors both within and outside of the firm.[1] Further, the tone at the top could extend to setting the compensation incentives of the tax director (who has direct involvement for the firm's tax decisions).[2] Anecdotal evidence is consistent with the tone at the top influence of top executives. For example, in deposition testimony, David Bullington, Wal-Mart's vice president for tax policy, stated that he began to feel pressure to lower the company's effective tax rate after the current chief financial officer, Thomas Schoewe, was hired in 2000. "Mr. Schoewe was familiar with some very sophisticated and aggressive tax planning ... And he rides herd on us all the time that we have the world's highest tax rate of any major company," Mr. Bullington said, according to a transcript of the deposition (Drucker 2007, A1).

It is important to clarify at the outset what we mean by the term "tax avoidance." As in Dyreng et al. (2008), we define tax avoidance broadly to encompass anything that reduces the firm's taxes relative to its pretax accounting income. We are not attempting to measure tax aggressiveness, tax risk, tax evasion, or tax sheltering. To keep our measures of tax avoidance broad and easy to understand, we examine two standard measures. The first is the firm's effective tax rate as defined under GAAP (hereafter, *GAAP ETR*), which is total tax expense (current plus deferred tax expense) divided by pre-tax accounting income (adjusted for special items). The second measure is the firm's cash taxes paid divided by pre-tax accounting income (adjusted for special items) (hereafter, cash effective tax rate, or *CASH ETR*). We use both measures in order to capture the various objectives managers may have when it comes to tax avoidance. To the extent that managers are concerned about reducing tax expense for financial accounting purposes, the use of our first measure, *GAAP ETR*, will capture the executive's propensity to affect this metric. To the extent that managers are concerned with reducing actual cash taxes paid, our second metric, *CASH ETR*, will capture the effects on this cash tax rate. We refer to the measures collectively as effective tax rates.[3]

---

[1] Anecdotal reports are that firms differ greatly in how involved tax advisors are in structuring transactions and that this variation likely contributes to differing levels of tax avoidance. For example, in a survey of 150 firms, CFO Research Services, in collaboration with Deloitte Tax LLP, report that roughly 50 percent of the firms surveyed said that the tax department has a substantial role in major transactions and only around 25 percent responded that the tax department played a substantial role in making operating decisions (CFO Research Services and Deloitte Tax LLP 2006).

[2] This idea is consistent with the discussion in Crocker and Slemrod (2005) and the findings in Armstrong et al. (2009) discussed below.

[3] Executives can avoid or manage the expense without affecting cash taxes paid and can affect cash taxes paid without affecting the total tax expense. For example, if firms (or executives) change their valuation allowance, tax contingency reserve, or the amount of foreign earnings designated as permanently reinvested they can change the *GAAP ETR* but these actions would not change the *CASH ETR*. Conversely, if a firm accelerates tax depreciation relative to book depreciation (which is an expense in the denominator–pre-tax income) it would reduce cash taxes paid but not the *GAAP ETR* (since the deferred tax expense would increase).

We use a technique developed in the economics literature by Bertrand and Schoar (2003) that involves tracking individual executives who served as top executives at more than one firm. Using this technique, we can observe whether particular executives appear to have systematic effects on their firms' tax avoidance. We can test, for example, whether Firm A's tax avoidance changes when Executive X is hired and whether it changes again when s/he leaves to run Firm B. The executives we track include the CEO, CFO, and other top executives listed in the ExecuComp database that were employed by more than one firm, which allows us to isolate executive effects apart from firm effects. We identify 908 executives who worked for at least two firms and were employed by each firm for at least three years between 1992 and 2006. Following the approach in Bertrand and Schoar (2003), we regress the firm's effective tax rate on firm fixed effects, year fixed effects, and executive fixed effects. Thus, all stationary characteristics of the firm are controlled in our specification through the firm fixed effect, and any time-specific, cross-sectional effects on effective tax rates are controlled through the year fixed effects. In our main tests we also control for time-varying characteristics of the firm through a set of control variables (e.g., size, leverage, R&D).

Results indicate that individual executives play a statistically and economically significant role in determining the level of tax avoidance that firms undertake. While the results are slightly stronger for CEOs (i.e., for executives who were CEOs in the last year of their career in our data), the results also hold for CFOs and other executives covered by ExecuComp. The frequency of significant executives is far greater than would be observed under the null hypothesis that tax avoidance is determined by the characteristics of the firm, but not the individual executive *per se*. We also examine executives in event-time as they move across firms, and we illustrate the effects they have on their firms' effective tax rates as they join and later depart from the firm. Finally, we conduct a number of robustness tests, which consistently show significant executive effects on both *GAAP ETR* and *CASH ETR*.

After establishing an executive effect, we investigate whether tax avoidance is associated with (1) other executive effects that represent broader styles such cost-cutting or growth strategies and/or (2) executive biographical characteristics. There appears to be little connection between these factors and tax avoidance. We note that the lack of a strong association with other styles or manager characteristics does not mitigate our main result that the executive matters. It simply means there are not identifiable, common characteristics that can explain executive-specific tax avoidance—the executive effects on tax avoidance appear to be idiosyncratic.

To our knowledge, ours is the first study to find and quantify the effects that individual top executives have on tax avoidance. As Bertrand and Schoar (2003) state, the neoclassical view of the firm is that the same decisions are made regardless of the executive in office. Prior and contemporaneous research in the tax avoidance area takes the neoclassical view, whereby tax avoidance is viewed as driven by firm characteristics and executive compensation but the individual executive is assumed to not matter.[4] Our study broadens the view of what drives tax avoidance to include the individuals in the top executive positions.

Our study is distinct from the prior and contemporaneous tax avoidance literature in three ways. First, we posit and find effects from the executives themselves. We view our consideration of the role of executives to be an important step toward a better understanding of the substantial variation in tax avoidance that prior research has found across firms, even among firms in the same industry (e.g., Dyreng et al. 2008). Second, because of our research design, we are able to quantify the tax avoidance effects of individual executives, something that no prior study has been able to do, and we find these effects are economically significant. Third, our results indicate that the

[4] Studies focusing on the role of compensation and tax avoidance include Phillips (2003), Armstrong et al. (2009), Rego and Wilson (2009), and Robinson et al. (2010).

executive is a significant determinant of *CASH ETR*, an important finding given that none of the aforementioned studies investigating firm-level incentives document any effect on *CASH ETR*.

Our findings raise many questions that we hope will be the addressed in future research. What happens to these executives and the firms they manage in the future? Tax avoidance can be value-enhancing to the firm but not necessarily so. Perhaps the executives who appear to emphasize tax avoidance simply make their firms take on substantial tax risk. In that case, we might expect to see additional taxes being paid in future years as the firms undergo audit by the IRS, perhaps on another executive's watch. It would be interesting to know what happens to the career prospects of executives who emphasize or de-emphasize tax avoidance. These are just a few of the questions raised by this study that we believe can lead to further research.

In the next section we review the prior literature, specifically the prior research on executive effects and the literature on tax avoidance. In Section III we discuss our sample selection and present descriptive statistics. We present the empirical analysis in Section IV and conclude in Section V.

II. PRIOR LITERATURE

Prior Literature on Tax Avoidance

Tax avoidance has been of interest to researchers for decades, yet surprisingly little empirical evidence exists about cross-sectional variation in tax avoidance, and until recently there has been little research on executives' roles in tax avoidance. We discuss both briefly here and refer readers to Hanlon and Heitzman (2010) for a broad review of the literature.

Firm-level characteristics such as size, economies of scale via foreign operations, tax planning, and other factors have been examined as determinants of tax avoidance, where tax avoidance has been measured in a variety of ways.[5] In a pioneering study that examines the role of manager incentives, Phillips (2003) uses private survey data to investigate whether compensating managers based on pre-tax or after-tax earnings affects *GAAP ETRs*. Phillips (2003) finds that compensation for managers of business units based on income measured after taxes (as compared to before tax) appear to be associated with a lower GAAP effective tax rate; however, this same association does not hold for CEO performance measures.[6] Recent studies in this area of incentive effects on tax avoidance include Desai and Dharmapala (2006), which examines "high-powered" incentives and tax avoidance; Armstrong et al. (2009), which is the first to examine compensation of tax directors; Rego and Wilson (2009), which examines links between top executive compensation and tax aggressiveness; and Robinson et al. (2010), which examines the effects on ETRs of treating the tax department as a profit center. Finally, in studies of tax shelter firms (an alternative measure of tax avoidance), Wilson (2009) and Lisowsky (2009) examine firm characteristics associated with tax sheltering.

There has been much less research on the consequences of tax avoidance. Desai and Dharmapala (2008) examine the effects of tax avoidance on firm value, and Hanlon and Slemrod (2009) examine market reactions to tax shelter involvement. Jennings et al. (2009) find an association between tax avoidance and implicit taxes; Graham and Tucker (2006) examine whether tax

---

[5] See Zimmerman (1983), Gupta and Newberry (1997), Mills et al. (1998), Rego (2003), Siegfried (1974), Porcano (1986), Stickney and McGee (1982), and Shevlin and Porter (1992). In addition, see Callihan (1994) for a summary of annual effective tax rate research.

[6] Phillips (2003, 869) interprets these results by stating that, "After-tax CEO performance measures could have an indirect negative effect on ETRs, however, because CEOs compensated after-tax are more likely to compensate their BU managers on an after-tax basis" and that other incentives such as job retention are perhaps sufficient to motivate CEOs to focus on after-tax results.

shelter firms carry less debt (i.e., whether the shelters are non-debt tax shields); and Wilson (2009) provides evidence that shelter firms have higher market returns as long as the firm's governance is "good" (measured by the Gompers et al. [2003] governance index).

In sum, prior research has investigated how to measure tax avoidance (e.g., Dyreng et al. 2008), the firm-level determinants of tax avoidance, and, to some extent, on the consequences of tax avoidance. While the effect of the CEO or other top management has recently been investigated with other performance outcomes of the firm (discussed in the next section), little attention has been given to whether individual executives have an impact on how much tax the firm avoids.[7] We address this question. This is an important step in identifying factors that make some firms more able to avoid tax than others. We do not investigate the consequences of tax avoidance resulting from executive effects. We look forward to research on this and related questions in the future.

The Effects of CEOs and Other Top Executives on Other Aspects of Firm Performance

How much of an effect the top executives and the "tone at the top" these executives set has on various aspects of firm performance and shareholder value is a question that is attracting increased attention. For example, Hayes and Schaefer (1999) estimate how much executives are worth to shareholders by identifying firm/executive separations and computing the abnormal returns surrounding the separations.[8] Bertrand and Schoar (2003) examine whether top executives affect firm performance and decisions by constructing a sample of executives who have moved across at least two firms and investigating the executives' effects on these firms. By tracking executives who change firms, we can separate the firm effect from the executive effect on corporate behavior and performance.

Several other studies examine the relation between CEO characteristics or CEO personal events and company performance. For example, Chatterjee and Hambrick (2007) examine narcissistic CEOs and company performance, while Liu and Yermack (2008) find that future company performance deteriorates when CEOs acquire very large homes or estates. Malmendier and Tate (2009) report evidence that CEOs who win awards from the business press (e.g., being named to *BusinessWeek*'s "Best Managers" list) underperform relative to their peer group after winning the award (similar to the *Sports Illustrated* "Jinx"). Bennedsen et al. (2008) use data from Denmark and report that CEOs' (but not board members') own and family deaths are strongly correlated with declines in firm operating performance. The most meaningful deaths for the CEOs, as might be expected, are the death of a child or a spouse, whereas the death of a mother-in-law has very little impact upon performance.

There are also several concurrent studies that use the Bertrand and Schoar (2003) methodology to examine executive effects on accounting outcomes. For example, Bamber et al. (2010) examine executive fixed effects on corporate financial disclosure policies and find that individual managers have a significant effect on voluntary disclosure (measured using management fore-

[7] There are other studies broadly related to the questions examined in this study. One is Chen et al. (2010), which examines whether family owned/run firms are more tax aggressive than non-family owned/run firms. The authors find that family owned/run firms are less aggressive because they try to convince minority shareholders that family members are not rent-extracting through complicated tax structures. Thus, while Chen et al. (2010) examine the effects of certain types of ownership it is not a test of whether executives matter to tax avoidance absent family ownership (note our study is about managers that change jobs so they are not likely members of the founding family). Another is Frank et al. (2009), which examines whether firms that are aggressive for tax purposes are aggressive for financial accounting purposes and whether firms that are aggressive for both have other similar corporate policies.

[8] See also Johnson et al. (1985), which examines the market reaction to the sudden deaths of executives (53 events). The authors find the market reaction is associated with founder status, decision-making responsibilities, and estimates of transaction costs for negotiating employment agreements. See also Pourciau (1993) and Murphy and Zimmerman (1993).

casts). Ge et al. (2009) investigate whether manager effects are significant in financial accounting outcomes such as discretionary accruals, off-balance-sheet activities, and conservatism.

III. SAMPLE AND DESCRIPTIVE STATISTICS

The sample begins with all executives listed in the ExecuComp database for the years 1992 to 2006. We track executives across firms and retain those executives who were employed for at least three years in each of at least two different firms. We impose minimal data requirements, requiring that the firm-year be listed on Compustat with total assets available. We then inspect each executive change to ensure that they were actual changes of firms. The resulting sample includes 12,958 firm-years of data, corresponding to 1,138 distinct firms, and 908 distinct executives.

Table 1 shows the breakdown of executives by their current title (i.e., the last title listed for the executive on ExecuComp) and their prior title. Overall, the sample contains 351 executives whose last title is CEO, 195 whose last title is CFO, and 362 whose last title is something else (e.g., president, vice-president). There are 101 instances where an executive was CEO of one firm listed in ExecuComp and moved to become CEO of another firm. There are few instances of a CEO of one firm going to take a lower title at another firm—in only four cases did a CEO later take on the title of CFO and in 58 cases a CEO later took on a non-CEO, non-CFO title. Being a CFO, on the other hand, is sometimes a stepping-stone to becoming CEO at another firm. This transition occurs 33 times in our sample, while 160 times a CFO makes a move to another firm but retains the same title. Finally, there are 217 instances of executives with titles other than CEO and CFO going on to become CEOs. In 31 cases a non-CEO/non-CFO moves to become a CFO and in 281 cases moves to another non-CEO/non-CFO position.

Table 2 presents descriptive statistics for the 12,958 firm-years in our sample. The effective tax rate variable (*GAAP ETR*) is measured as total tax expense (current and deferred) divided by

TABLE 1

Frequency of Executive Changes

	Current Title			
Prior Title	**CEO**	**CFO**	**Other**	**Total**
CEO	101	4	58	163
CFO	33	160	23	216
Other	217	31	281	529
Total	351	195	362	908

This table describes title changes as executives move from one company to another. The sample is a total of 908 executives (351 CEOs, 195 CFOs, and 362 other executives). Each executive is required to be employed by at least two different firms, for at least three years at each firm. The columns across the top describe the executives' title in the last firm in which the executives were observed (current title), while the rows describe the title at the firm in which the executives were employed before the last firm. For example, the cell in the first column and first row describes CEOs that are CEO in their current firm (last firm listed in ExecuComp) and were CEO at the firm where they were previously employed. In the case of executives who were employed by more than two firms for at least three years, the table represents their last change only.

TABLE 2
Descriptive Statistics

Variable	n	Mean	Std. Dev.	25th Percentile	50th Percentile	75th Percentile
GAAP ETR	10,355	0.309	0.154	0.249	0.337	0.381
CASH ETR	10,124	0.263	0.195	0.133	0.249	0.347
EBITDA	12,716	0.153	0.119	0.084	0.140	0.210
R&D	12,958	0.045	0.117	0.000	0.000	0.032
ADVERTISING	12,958	0.011	0.110	0.000	0.000	0.007
SG&A	9,545	0.297	0.254	0.136	0.241	0.381
CAPITAL EXPENDITURES	10,109	0.298	0.349	0.120	0.199	0.338
PERCENTAGE CHANGE IN SALES	12,926	0.155	0.350	0.004	0.087	0.215
LEVERAGE	12,586	0.246	0.186	0.097	0.237	0.357
CASH HOLDINGS	12,578	0.158	0.257	0.019	0.059	0.187
FOREIGN OPERATIONS	12,958	0.470	0.499	0.000	0.000	1.000
NET OPERATING LOSS	12,958	0.273	0.446	0.000	0.000	1.000
SIZE	12,958	7.699	1.819	6.424	7.587	8.878
ESTIMATED OPTION EXPENSE	12,958	0.028	0.132	0.000	0.000	0.005
INTANGIBLE TO TOTAL ASSETS	12,958	0.126	0.192	0.000	0.038	0.182
GROSS PP&E TO TOTAL ASSETS	12,089	0.643	0.453	0.283	0.548	0.947

This table describes the variables used in the study. As each variable is described, the Compustat pneumonic is in parentheses.

Variable Definitions:

GAAP ETR = the financial accounting effective tax rate, defined as total income tax expense (*TXT*) divided by pre-tax book income (*PI*) before special items (*SPI*);

CASH ETR = the cash effective tax rate, defined as cash tax paid (*TXPD*) divided by pre-tax book income (*PI*) before special items (*SPI*);

EBITDA = earnings before interest, taxes, depreciation, and amortization (*OIBDP*) scaled by lagged total assets (*AT*);

R&D = research and development expense (*XRD*) divided by net sales (*SALE*); when missing, reset to 0;

ADVERTISING = advertising expense (*XAD*) divided by net sales (*SALE*); when missing, reset to 0;

SG&A = selling, general, and administrative expense (*XSGA*) divided by net sales (*SALE*); missing values of *SG&A* are set to 0 when used as a control variable in Equation (1);

CAPITAL EXPENDITURES = reported capital expenditures (*CAPX*) divided by gross property, plant, and equipment (*PPEGT*);

PERCENTAGE CHANGE IN SALES = the annual percentage change in net sales $((SALE_t / SALE_{t-1}) - 1)$;

LEVERAGE = the sum of long-term debt (*DLTT*) and long-term debt in current liabilities (*DLC*) divided by total assets (*AT*);

CASH HOLDINGS = cash and cash equivalents (*CHE*) divided by total assets (*AT*);

FOREIGN OPERATIONS = the firm has a non-missing, non-zero value for pre-tax income from foreign operations (*PIFO*);

SIZE = the natural log of total assets (*AT*);

NET OPERATING LOSS = an indicator if the firm has a non-missing value of tax loss carry-forward (*TLCF*);

ESTIMATED OPTION EXPENSE = calculated from ExecuComp as the average annual value realized from exercise of options for the top executives grossed up by the fraction of options owned by the covered executives, scaled by average total assets;

INTANGIBLE TO TOTAL ASSETS = the ratio of intangible assets (*INTANG*) to total assets (*AT*); and

GROSS PP&E TO TOTAL ASSETS = gross property, plant, and equipment (*PPEGT*) divided by total assets (*AT*).

pre-tax income.[9] The mean *GAAP ETR* in the sample is 30.9 percent, with a median of 33.7 percent. We also measure firms' cash effective tax rates (*CASH ETR*) as the ratio of cash taxes paid to pretax income. The sample firms have *CASH ETR*s with a mean of 26.3 percent and a median of 24.9 percent.[10] The observation that *CASH ETR* is lower than *GAAP ETR* is consistent with firms on average having lower taxable income than pretax accounting income, and is consistent with prior research. In addition, we recognize that *CASH ETR* is more volatile than *GAAP ETR* and potentially noisy on an annual basis and, thus, results may be more difficult to document (Dyreng et al. 2008). Note that the number of observations changes over the measures because if the denominator is negative for either ratio or if the cash taxes paid is not disclosed or negative, then we code that observation as missing.

We also present a number of other variables in Table 2 that are used later when we investigate whether executives' tax avoidance effects are associated with the executives' effects on other corporate actions and outcomes. These other nontax variables are all defined in Table 2 and include measures such as *EBITDA*, research and development expense (*R&D*), advertising expense (*ADVERTISING*), and selling, general, and administrative expense (*SG&A*) (unless otherwise noted, variables are scaled by beginning total assets). Because we include firms that have a departing/incoming executive at some point in the sample period, we compare our descriptive statistics to those of the Compustat population to get a sense of whether our firms are somehow different, which would limit generalizability. We find that our sample firms are slightly more profitable and invest slightly more in R&D, but are not substantially different from the typical firm on Compustat (results untabulated). We next turn to the primary empirical analysis.

IV. EMPIRICAL ANALYSIS

Primary Specification

Our primary research question is whether individual executives matter for firm tax avoidance. As discussed in Bertrand and Schoar (2003), much of the prior literature has either assumed that all executives are homogeneous and provide selfless inputs into the firm's processes or that even if executives are heterogeneous, they cannot easily affect firm policies. An opposing view from standard agency models is that executives have discretion within their own firms and use this discretion to affect corporate decision making and possibly to advance their own objectives.

The identification strategy is a regression based on Bertrand and Schoar (2003) that uses a panel of data:

$$ETR_{it} = \alpha_0 + \sum_k \alpha_k CONTROL_{it}^k + \sum_t \alpha_t YEAR_t + \sum_i \alpha_i FIRM_i + \sum_m \alpha_m EXEC_m + \varepsilon_{it}, \quad (1)$$

where ETR_{it} is *GAAP ETR* (or *CASH ETR*, each tested in separate regressions); $FIRM_i$ is a separate indicator variable for each firm, i (firm fixed effects); $YEAR_t$ is a separate indicator variable for each year, t (year fixed effects); ε_{it} is the error term. $EXEC_m$ is a separate indicator for each executive, m (executive fixed effects), and is our main variable of interest. The firm fixed effects in the model control for all stationary firm characteristics such as industry and stationary firm-level strategies for tax avoidance. The vector of control variables, $CONTROL^k$, reflects firm

[9] Pretax income is measured as income before discontinued operations and extraordinary items and excludes special items. ETRs with negative pretax income are set to missing. The remaining non-missing ETRs are winsorized (reset) so that the largest observation is 1 and the smallest is 0. While we would like to measure long-run cash effective tax rates as in Dyreng et al. (2008), it is not possible for us to examine both manager changes and estimate tax rates over a ten-year span.

[10] The annual cash tax rates are similar to the rates in Dyreng et al. (2008), which used a larger sample but required firms to have positive income summed over a ten-year period. The mean annual cash tax rate (measured similarly) in Dyreng et al. (2008) is 27 percent and the median is 25.6 percent.

characteristics that can change over time and could plausibly be related to tax avoidance, as is discussed in more detail below.

We retain firm-years with an executive change and all available prior and subsequent years for the firm. The years outside the executive tenure allow a better estimation of the firm effect versus the executive effect. Thus, we employ a panel data set in all tests.

Within the title *EXEC*, we also separately identify executives by title (e.g., CEO, CFO, or some non-CEO, non-CFO title). To do this, we classify executives according to the title last listed on ExecuComp for the individual. Coding each title separately allows us to test each group of executives while controlling for the effects of the other executives (for example, using a joint F-test, we can test the effect of CEOs while controlling for the effects of CFOs).

If we find that executives matter, it could be that executives have different styles that they impose on the firm or it could be that the firm hired the executives because of their style (e.g., a tax avoider). Because we control for the firm effect, however, only if the firm changed objectives and hired executives with different styles through time could our results be due to matching between the firm and executives; otherwise the new executive's style would be the same as the old executive's and the effect would be captured by the firm indicator variable.

As noted earlier, $CONTROL^{K}$ in the above equation represents a vector of time-varying, firm-level controls. When deciding upon a research design, we gave considerable thought to the variables to include in $CONTROL^{K}$. The more variables we include in $CONTROL^{K}$, the less variation we leave for the executive effect to "pick up." In other words, if executive choices affect the variables in our vector of controls, then including these controls could drive the executive fixed effect coefficients toward zero. For example, consider an executive's choice to increase spending on research and development (R&D). Assume that this spending generates additional research and development tax credit, which reduces the effective tax rates of the firm (note that R&D spending is expensed for financial reporting so the denominator of our ratios is also reduced, but assuming the amount is creditable likely results in reduced effective rates). What we cannot separate in our tests is whether the executive increased the R&D spending in order to reduce the tax rate or whether the executive increased the R&D spending for other reasons and a byproduct of this decision is a reduced tax rate. If we include R&D as a control variable in the regression, then the implicit assumption in the research design is that the R&D spending was for a nontax reason that we want to control for in our tests and remove from the manager effect coefficient. If we do not include R&D as a control variable in the regression, then the implicit assumption in the research design is that the effect of the increased R&D spending on the tax rate is intentional on the part of management, such that the regression will place that effect in the executive's fixed effect coefficient. Since we obviously do not know the extent to which each decision an executive makes is explicitly for tax reasons, we have to arbitrarily choose between these two designs and interpret the results accordingly.

With this caveat in mind, we have chosen to estimate our main analyses after including a wide range of variables in $CONTROL^{K}$. We draw these variables from prior effective tax rate literature (e.g., Mills et al. 1998; Rego 2003). We include *ADVERTISING* expense, *CAPITAL EXPENDITURES*, *ESTIMATED OPTION EXPENSE*, an indicator for whether the firm has *FOREIGN OPERATIONS*, the ratio of *INTANGIBLE TO TOTAL ASSETS*, *LEVERAGE*, *SIZE*, an indicator for whether the firm has a *NET OPERATING LOSS (NOL)*, *GROSS PROPERTY PLANT & EQUIPMENT*, *SG&A* expense and *R&D* expense.[11] We also estimate the regression with only a minimal set of controls (*NOL* and *SIZE*) in an untabulated robustness check, finding very similar results.

[11] Variable definitions are found in Table 2. When using SG&A as a control variable in Equation (1), missing values are set to 0.

It is again important to note that by including firm fixed effects, we control for stationary firm characteristics. For example, if Firm A has lower tax rates than Firm B because Firm A always has more foreign operations than Firm B, then that effect will be captured in the fixed effect for Firm A. Similarly, the tests control for industry membership because industry is a stationary firm characteristic. However, changes in Firm A's effective tax rates caused by a new executive who changes Firm A's foreign operations will be reflected in the fixed effect for that executive. Thus, the intention is for the executive fixed effect to capture the impact on tax avoidance from changes at the firm while under his/her command.

Our sample consists of executive changes. If a firm had the same managerial team over time, then the firm fixed effects and executive fixed effects would be perfectly collinear and we could not separate the effects that are specific to an executive from the effects of the firm. The variation in executives over time (i.e., executive changes) is what makes it possible to separate executive effects from firm effects. Moreover, by observing the performance of the same executive across two firms, say A and B, we reduce the likelihood that the results are due to coincidental time-varying performance in one of the firms. Accordingly, we impose the same requirement on our sample as did Bertrand and Schoar (2003) by requiring executives in our sample to have been employed by at least two firms in the sample, with at least three years of employment at each firm.[12]

Main Results

Table 3 reports the results from estimating Equation (1) across the sample of firm-years with non-missing tax rates (n = 10,355 for *GAAP ETR* and n = 10,124 for *CASH ETR*). Panel A presents the results with *GAAP ETR* as the dependent variable and Panel B with *CASH ETR* as the dependent variable. The number presented in the table is the F-statistic from the joint test of the significance of that particular set of coefficients. (For example, 10.00 in Panel A is the F-statistic from the joint test of significance of the coefficients on the year indicator variables.)

Model 1 is a baseline regression that includes only an intercept and the vector of time-varying firm controls ($CONTROL^{K}$) described above. Moving to the right, we present the results from a model that includes the vector of time varying controls and only year effects (Model 2), only firm effects (Model 3), only CEO effects (Model 4), only CFO effects (Model 5), and only other (non-CEO/non-CFO) executive effects (Model 6). Model 7 presents results of a regression that only includes the time varying controls ($CONTROL^{K}$) and only the executive fixed effects. In this model, we include an indicator variable for all executives in the sample and test them jointly, regardless of specific title (i.e., CEO, CFO, Other). In Model 8, we estimate the full model specified in Equation (1). Here we include the executive fixed effect (indicator variable) for every executive as well as the firm fixed effect and year fixed effect and the time varying controls. As described above, when specified in this manner, we can statistically test the significance of any subset of indicator variables. We report tests of each subset of executives in Model 8. The difference across models when looking at the F-statistic for CEOs, for example, is that in Model 8, the F-statistic tests the joint significance of all the CEOs after controlling for time varying controls,

[12] It is possible that teams of executives move from firm to firm, but not in exactly the same years. We find this phenomenon occurs in about 10 percent of cases. We examine the influence of this phenomenon on the executive's ability to affect the firm's *GAAP ETR* or *CASH ETR* in Table 7. We note we cannot test the effect of a team moving in the exact years as we cannot separate out the executive effect from the team effect and we cannot associate this team effect with any executive characteristics such as M.B.A., age, etc., in our later tests below. Thus, these executives where the team moves in the exact same years are removed from our sample (i.e., the perfectly collinear executives). There are 24 such executives.

TABLE 3

Test of Manager Effects

Panel A: Test of Manager Effect on Effective Tax Rates (*GAAP ETR*)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
YEAR		10.00***						12.01***
FIRM			2.62***					2.39***
CEO				2.07***				1.66***
CFO					1.92***			1.70***
OTHER						1.95***		1.63***
EXEC							2.01***	1.75***
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES
n	10,355	10,355	10,355	10,355	10,355	10,355	10,355	10,355
NYEARS		14						14
NFIRMS			1,007					1,007
NCEOS				311				311
NCFOS					172			172
NOTHERS						312		312
NEXECS							795	795
ADJRSQ	0.030	0.041	0.162	0.060	0.044	0.057	0.100	0.226

Panel B: Test of Manager Effect on Cash Effective Tax Rates (*CASH ETR*)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
YEAR		16.10***						14.14***
FIRM			2.95***					2.49***
CEO				2.15***				1.67***
CFO					1.87***			1.36***
OTHER						1.96***		1.61***
EXEC							2.00***	1.57***
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES
n	10,124	10,124	10,124	10,124	10,124	10,124	10,124	10,124
NYEARS		14						14

(continued on next page)

Panel B: Test of Manager Effect on Cash Effective Tax Rates (*CASH ETR*)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
NFIRMS			996					996
NCEOS				307				307
NCFOS					172			172
NOTHERS						304		304
NEXECS							783	783
ADJRSQ	0.024	0.044	0.181	0.057	0.038	0.052	0.094	0.236

*, **, *** Represent statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

This table presents F-statistics from various specifications of Equation (1) as follows:

$$ETR_{it} = \alpha_0 + \sum_k \alpha_k CONTROL_{it}^k + \sum_t \alpha_t YEAR_t + \sum_i \alpha_i FIRM_i + \sum_m \alpha_m EXEC_m + \varepsilon_{it}$$

Each cell in Panels A and B of the table is an F-statistic testing the significance of the effects listed in the first column and each column represents a regression nested within the model. For example, the number 10.00 in the column labeled Model 2 of Panel A is the F-statistic from the test of whether the year fixed effects are significant in a regression where *GAAP ETR* is the dependent variable. Model 1 includes only the vector of time varying controls (*ADVERTISING*, *CAPITAL EXPENDITURES*, *ESTIMATED OPTION EXPENSE*, *FOREIGN OPERATIONS*, *INTANGIBLE TO TOTAL ASSETS*, *LEVERAGE*, *SIZE*, *NET OPERATING LOSS*, *GROSS PROPERTY PLANT & EQUIPMENT*, *SG&A*, and *R&D*). Missing values of *ADVERTISING*, *SG&A*, and *R&D* are set to 0 when used as a control variable in Equation (1). Models 2 through 6 are regressions, which include only one set of effects (indicator variables)—*YEAR*, or *FIRM*, or *CEO*, etc., but include no controls for the other effects (except the time varying effects, which are included in all models). Model 7 includes indicators for all executives regardless of title and the F-statistic listed above (2.01 in Panel A and 2.00 in Panel B) is the F-statistic on the joint test of whether all those indicator variables are significant. The final model, Model 8, is a test of each set of effects in the presence of the other effects. In this model an indicator for each executive is included, and we test the groups of executives separately (e.g., 1.66 F-statistic in Panel A for the effect of the CEO after controlling for all other effects) and collectively (e.g. 1.75 F-statistic for the combined effect of all executives regardless of title; Bertrand and Schoar 2003).

the firm effect, the year effect, and the effects of executives with other titles, whereas in Model 4, the F-statistic tests the joint significance of all the CEOs after controlling for the time varying effects only.

Results indicate that both firm and year fixed effects are significant when included on their own in both the *GAAP ETR* and the *CASH ETR* regressions (Models 2 and 3) in Panels A and B. All three sets of executive coefficients are significant at the 1 percent level when considered individually (Models 4–6) and collectively (Model 7). Model 8 presents the results for the full model from Equation (1) that includes the vector of time varying controls (*NOL*, *SIZE*, *R&D*, *SG&A*, etc.), fixed effects for the firm, the year, and each executive. Each type of executive has a fixed effect that is significant in the full model in both the *GAAP ETR* and the *CASH ETR* regressions (i.e., after controlling for the effect of other executives and firm and year effects, each type of executive still has a significant fixed effect, on average). The F-statistics are usually largest for CEO fixed effects, both in the full model and when each executive group is tested individually, which might be expected if CEOs (and future CEOs) have the most influence on the firm's policies. However, we consistently find significant effects for non-CEO executives as well.

These results document that executives have a significant effect on *GAAP ETR*s and on *CASH ETR*s. This effect is after controlling for many of the firm-level items that have been documented (or suspected) to affect these effective tax rates, such as R&D spending, foreign operations, and other variables. This is an important step in the literature, as heretofore it has been difficult to find factors that are consistent determinants of effective tax rates (even the effect of *SIZE* is debated in the literature). In this study, however, we document an important item not previously considered—executives have a significant effect on the level of effective tax rates of the firm.

Robustness Tests

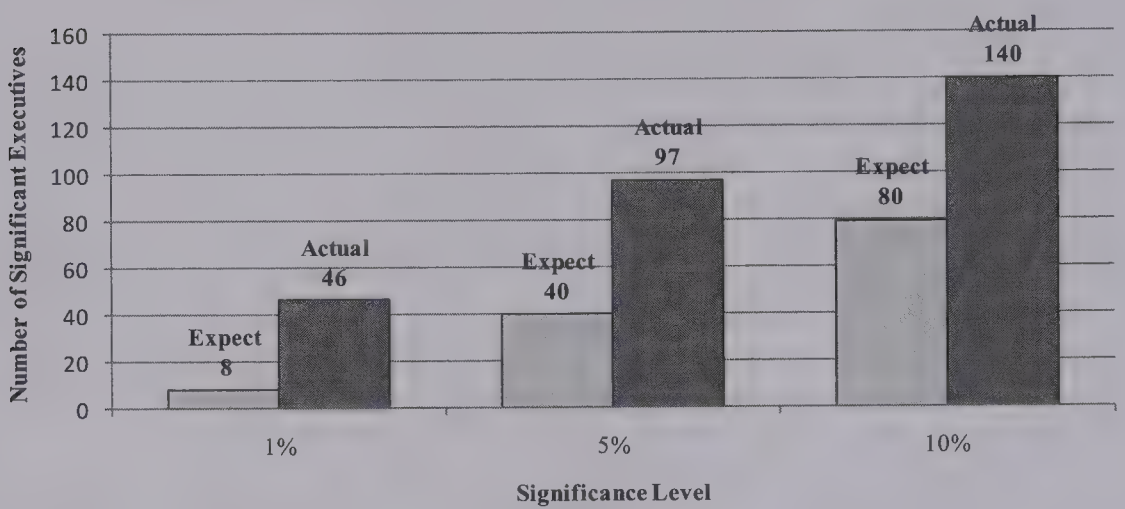
We conduct a battery of robustness tests on our main results. In the first set of tests, we examine the effect of the time varying control variables on the significance of the executive fixed effects. In untabulated tests, we re-estimate Equation (1), but reduce the vector of time varying controls to include only *NET OPERATING LOSS* and *SIZE*. We find results are very similar to those in Table 3.

In the second set of robustness tests, we address the potential concern that our regressions may show significance for an executive even if the reduction to the effective tax rates was due to some random event that occurred during his/her tenure. In order to mitigate this concern, we examine whether the significant executive effects are driven by just one of the firms at which the executive works or whether executive effects persist across the firms at which the executive is employed. We do this by estimating a model similar to Model 8 in Table 3, but without the executive fixed effects, thus leaving the executive effects in the residual. We then regress each significant (at the 5 percent level) executive's mean residual corresponding to his/her tenure from his/her second firm on the mean residual corresponding to his/her tenure at his/her first firm. In untabulated results, we find that the tax avoidance at the first firm is positively associated with the tax avoidance at the second firm; the coefficient is 0.353 ($p = 0.023$) for the *GAAP ETR* regression and 0.280 ($p = 0.053$) for the *CASH ETR* regression. Thus, the results do not appear to be driven by an effect at only one of the firms at which the executive is employed.

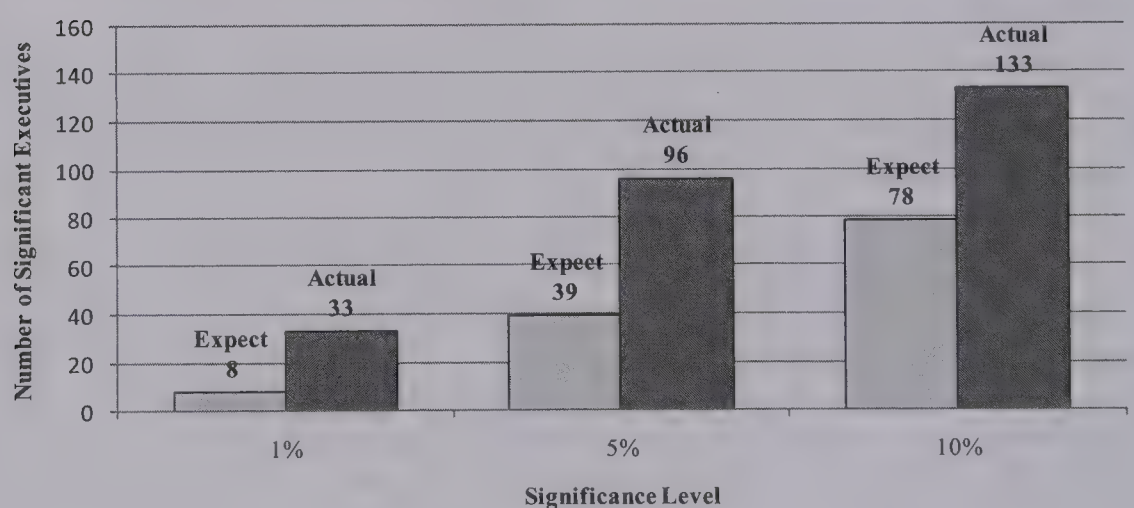
In the third set of tests, we consider alternatives to the tests based on the F-statistics, since it is possible for F-statistics to be influenced by only a few significant coefficients. Accordingly, we examine the frequency of executives whose specific fixed effect is significant (i.e., the t-statistic on the fixed-effect coefficient is significant). Figure 1 presents the number of individual executives with statistically significant fixed executive effects from Model 8 in Table 3. Our sample contains 795 (783) executives with sufficient data to estimate the *GAAP ETR* (*CASH ETR*) regression. Under the null hypothesis that executives have no incremental effect beyond firm fixed effects and

FIGURE 1
Frequencies of Significant Manager Fixed Effects

Panel A: Frequency of Significant Fixed Effects Compared to Expected Frequency with *GAAP ETR* as Variable of Interest



Panel B: Frequency of Significant Fixed Effects Compared to Expected Frequency with *CASH ETR* as Variable of Interest



Panels A and B of the figure report the frequency of significant executive fixed effects relative to what would be expected under the null hypothesis if the models tested in Table 3 are well specified. The calculations are performed at the 1 percent, 5 percent, and 10 percent significance level, respectively. Panel A presents the calculation when *GAAP ETR* is the dependent variable (as in Table 3 Panel A), and Panel B presents the calculation when *CASH ETR* is the dependent variable (as in Table 3 Panel B).

time varying characteristics, we would expect 80 executives to be significant in the *GAAP ETR* regression at the 10 percent level (795 × 0.10), 40 to be significant at the 5 percent level, and 8

to be significant at the 1 percent level of significance. In our regression results, when using *GAAP ETR* as the dependent variable (Panel A in Table 3), 140 executives are significant at the 10 percent level, 97 are significant at the 5 percent level, and 46 are significant at the 1 percent level. When using *CASH ETR* as the dependent variable (Panel B of Table 3), 133 executives are significant at the 10 percent level, 96 are significant at the 5 percent level, and 33 are significant at the 1 percent level. Thus, we find the number of significant executives in our data to be consistently greater than what we would expect under the null. In untabulated data, the results for each separate level of executive (e.g., *CEO*, *CFO*, *OTHER*) are similar—the actual number of significant executives is much greater than one would expect under the null hypothesis.

In the fourth set of robustness tests, we control for the possibility that the standard errors on the t-statistics and F-statistics could be downwardly biased for any reason. Accordingly, we re-conduct the tests in Table 3, but instead of actual data, we use data that have been randomly shuffled. Specifically, we randomize the data to generate the empirical distribution of estimated executive effects under the null hypothesis by shuffling the executive-firm pairings so that each executive is assigned to a firm at which she/he did not actually work. We repeat this shuffling 1,000 times and compute the percentage of significant executive effects for each iteration. Because we would not expect a randomly matched firm and executive to generate significant results, this procedure generates the empirical distribution under the null hypothesis against which our actual results can be compared.

The results in both the *GAAP ETR* and *CASH ETR* regressions reveal that more executives are significant in the actual data than *any* of the 1,000 iterations of the randomized data. Thus, the results are highly significant.

In the fifth set of robustness tests, we cluster standard errors at the firm level (Petersen 2009). However, clustering of the standard errors does not allow for the inclusion of all of our currently included fixed effects.[13] Thus, in order to estimate the regression we must eliminate observations (specifically those in which the executive is included in a cluster of firms where the number of executives within the cluster is equal to or exceeds the number of firms).[14] The results on this reduced sample are robust to using clustered standard errors. In fact, in all cases, the test statistics are stronger than those reported.

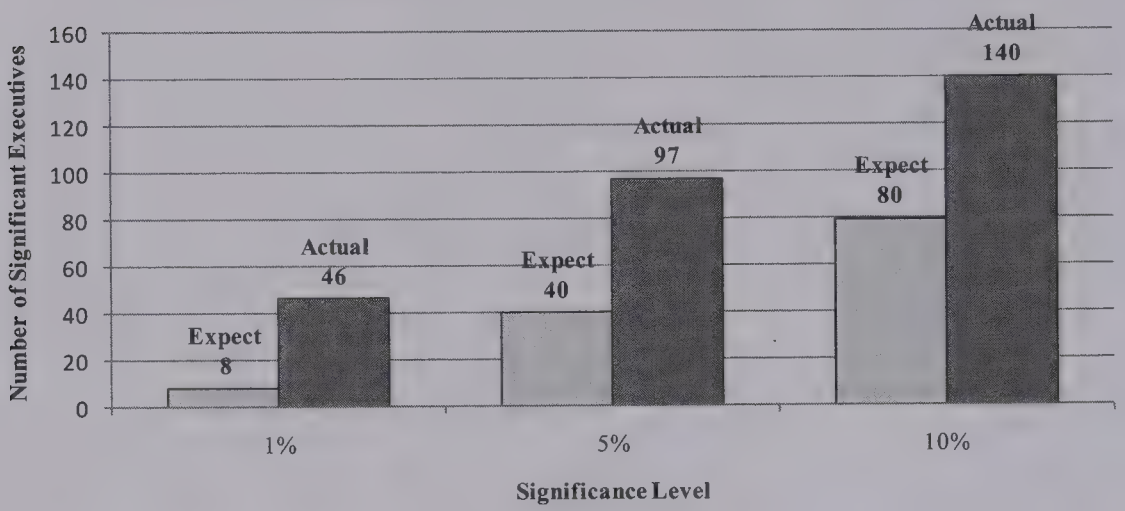
In the sixth robustness test, we recognize that variation in the ratios we examine, *GAAP ETR* and *CASH ETR*, could be driven by the either the numerator (income tax expense or cash tax paid) or by the denominator (pre-tax income). To test whether the executive effects on the *GAAP ETR* and *CASH ETR* ratios are being driven by executives who affect the denominator via earnings management, we change the dependent variables by putting cash from operations (adjusted for cash taxes paid) in the denominator. This change greatly reduces the possibility that variation in the ratio is driven by earnings management, although the interpretation of the ratios becomes less intuitive. In an unreported analysis similar to that in Table 3, we find that in all cases executive fixed effects are significant when we conduct the tests with cash flow from operations in the denominator. Thus, it appears that executives are able to influence the effective tax rate ratios via management of tax expense and/or cash tax paid, and not just by managing the denominator of the

[13] More specifically, clustering standard errors at the firm level restricts the effective degrees of freedom for the model to the number of clusters. Because we include firm fixed effects, we would be estimating a model where the number of parameters would exceed the effective degrees of freedom. While it is possible to generate standard errors for each of the variables in the regression, it is not possible to test more linear restrictions than degrees of freedom. Thus, in our particular setting it is not possible generate the test statistics associated with the overall model fit or the joint significance of the firm effects—which is our main test.

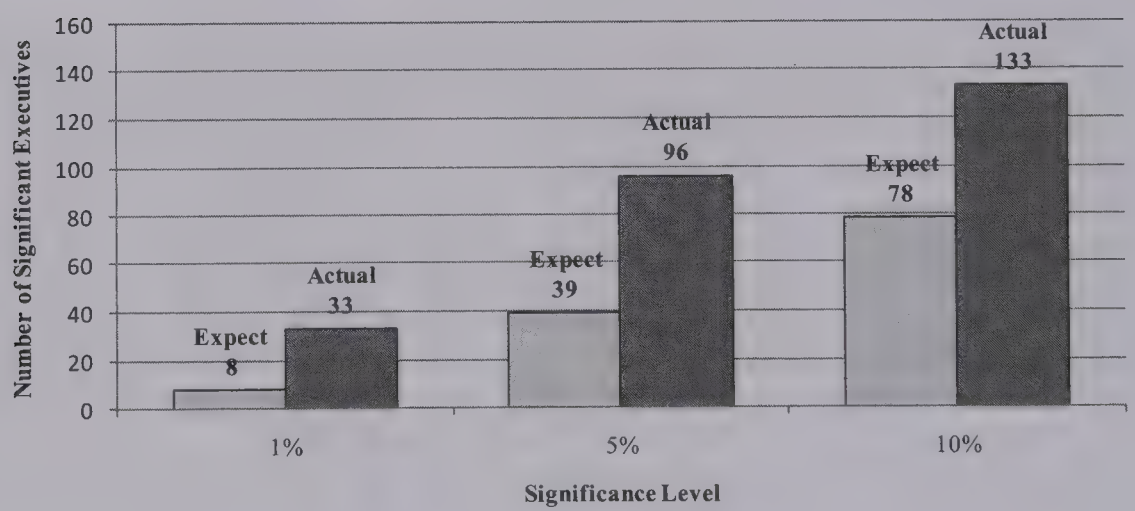
[14] This situation can arise if Executive X worked at Firm A and Firm C, Executive Y worked at Firm A and Firm B, and Executive Z worked at Firm B and Firm C. This results in three executives, but only three degrees of freedom, because there are only three firms.

FIGURE 1
Frequencies of Significant Manager Fixed Effects

Panel A: Frequency of Significant Fixed Effects Compared to Expected Frequency with *GAAP ETR* as Variable of Interest



Panel B: Frequency of Significant Fixed Effects Compared to Expected Frequency with *CASH ETR* as Variable of Interest



Panels A and B of the figure report the frequency of significant executive fixed effects relative to what would be expected under the null hypothesis if the models tested in Table 3 are well specified. The calculations are performed at the 1 percent, 5 percent, and 10 percent significance level, respectively. Panel A presents the calculation when *GAAP ETR* is the dependent variable (as in Table 3 Panel A), and Panel B presents the calculation when *CASH ETR* is the dependent variable (as in Table 3 Panel B).

time varying characteristics, we would expect 80 executives to be significant in the *GAAP ETR* regression at the 10 percent level (795 × 0.10), 40 to be significant at the 5 percent level, and 8

to be significant at the 1 percent level of significance. In our regression results, when using *GAAP ETR* as the dependent variable (Panel A in Table 3), 140 executives are significant at the 10 percent level, 97 are significant at the 5 percent level, and 46 are significant at the 1 percent level. When using *CASH ETR* as the dependent variable (Panel B of Table 3), 133 executives are significant at the 10 percent level, 96 are significant at the 5 percent level, and 33 are significant at the 1 percent level. Thus, we find the number of significant executives in our data to be consistently greater than what we would expect under the null. In untabulated data, the results for each separate level of executive (e.g., *CEO*, *CFO*, *OTHER*) are similar—the actual number of significant executives is much greater than one would expect under the null hypothesis.

In the fourth set of robustness tests, we control for the possibility that the standard errors on the t-statistics and F-statistics could be downwardly biased for any reason. Accordingly, we re-conduct the tests in Table 3, but instead of actual data, we use data that have been randomly shuffled. Specifically, we randomize the data to generate the empirical distribution of estimated executive effects under the null hypothesis by shuffling the executive-firm pairings so that each executive is assigned to a firm at which she/he did not actually work. We repeat this shuffling 1,000 times and compute the percentage of significant executive effects for each iteration. Because we would not expect a randomly matched firm and executive to generate significant results, this procedure generates the empirical distribution under the null hypothesis against which our actual results can be compared.

The results in both the *GAAP ETR* and *CASH ETR* regressions reveal that more executives are significant in the actual data than *any* of the 1,000 iterations of the randomized data. Thus, the results are highly significant.

In the fifth set of robustness tests, we cluster standard errors at the firm level (Petersen 2009). However, clustering of the standard errors does not allow for the inclusion of all of our currently included fixed effects.[13] Thus, in order to estimate the regression we must eliminate observations (specifically those in which the executive is included in a cluster of firms where the number of executives within the cluster is equal to or exceeds the number of firms).[14] The results on this reduced sample are robust to using clustered standard errors. In fact, in all cases, the test statistics are stronger than those reported.

In the sixth robustness test, we recognize that variation in the ratios we examine, *GAAP ETR* and *CASH ETR*, could be driven by the either the numerator (income tax expense or cash tax paid) or by the denominator (pre-tax income). To test whether the executive effects on the *GAAP ETR* and *CASH ETR* ratios are being driven by executives who affect the denominator via earnings management, we change the dependent variables by putting cash from operations (adjusted for cash taxes paid) in the denominator. This change greatly reduces the possibility that variation in the ratio is driven by earnings management, although the interpretation of the ratios becomes less intuitive. In an unreported analysis similar to that in Table 3, we find that in all cases executive fixed effects are significant when we conduct the tests with cash flow from operations in the denominator. Thus, it appears that executives are able to influence the effective tax rate ratios via management of tax expense and/or cash tax paid, and not just by managing the denominator of the

[13] More specifically, clustering standard errors at the firm level restricts the effective degrees of freedom for the model to the number of clusters. Because we include firm fixed effects, we would be estimating a model where the number of parameters would exceed the effective degrees of freedom. While it is possible to generate standard errors for each of the variables in the regression, it is not possible to test more linear restrictions than degrees of freedom. Thus, in our particular setting it is not possible generate the test statistics associated with the overall model fit or the joint significance of the firm effects—which is our main test.

[14] This situation can arise if Executive X worked at Firm A and Firm C, Executive Y worked at Firm A and Firm B, and Executive Z worked at Firm B and Firm C. This results in three executives, but only three degrees of freedom, because there are only three firms.

ratio. In sum, the weight of the evidence throughout the tests indicates that executives are important factors in the extent to which firms undertake tax avoidance.

The remaining robustness tests examine whether industry effects could be driving our results. First, we examine whether the significant executives are clustered in a few industries or are spread out over a variety of industries. We find that they are well spread across industries. For example, in untabulated results using the 17-industry Fama-French grouping, we find significant managers in 15 out of 17 industries using *GAAP ETR* and 17 of 17 industries using *CASH ETR*.[15] Moreover, the percentage of significant executives closely tracks the percentage of observations in each industry. Similar findings obtain when we use 30 industry groupings.

Second, we separately examine the distribution of significant positive and negative executive coefficients across industries. We again find significant executives to be well distributed across industries. Third, we examine results separately for executives who switch industries (43 percent of executives) and do not switch industries (57 percent of executives) when they switch jobs. As in the other tests, the significant executives are well spread across industries for both groups. Overall, the results do not appear to be driven by industry effects.

Event Time Analysis of Effective Tax Rates

We next illustrate the changes in *GAAP ETR* and *CASH ETR* in the periods before and after the executives join their respective firms and the periods before and after the executives depart their firms. We begin by classifying executives into one of four categories based on our regression results above: *GAAP ETR*-reducing executives (those with a significantly negative fixed effect at the 5 percent level in the Table 3, Model 8 regression), *GAAP ETR*-increasing executives (those with a significantly positive fixed effect at the 5 percent level), no effect *GAAP ETR* executives with positive coefficients (those with an insignificant positive fixed effect), and no effect *GAAP ETR* executives with negative coefficients (those with an insignificant negative fixed effect). We then normalize each firm's string of *GAAP ETR* data by the mean *GAAP ETR* in the three years leading up to the executive joining the firm ($t-3$ to $t-1$). This procedure allows us to interpret the figure as the change in the firm's *GAAP ETR* incremental to what the firm's average *GAAP ETR* was before the executive arrived. Because the data are classified based upon our regression results, we do not view these as independent tests, but rather as a graphical illustration of what a significant executive fixed effect indicates is occurring in the data.

Figure 2, Panel A shows the change in firms' *GAAP ETR* around the date of hiring. Firms run by *GAAP ETR*-increasing executives see their *GAAP ETR* increase by approximately 15 percentage points by three years after the year of hiring. Firms run by *GAAP ETR*-decreasing executives see their *GAAP ETR* decrease steadily by slightly more than 15 percentage points by year $t+3$. Firms run by executives whose fixed effects are not significant in Table 3 show changes in their *GAAP ETRs* of about five percentage points in absolute value after the executive joins the firm.

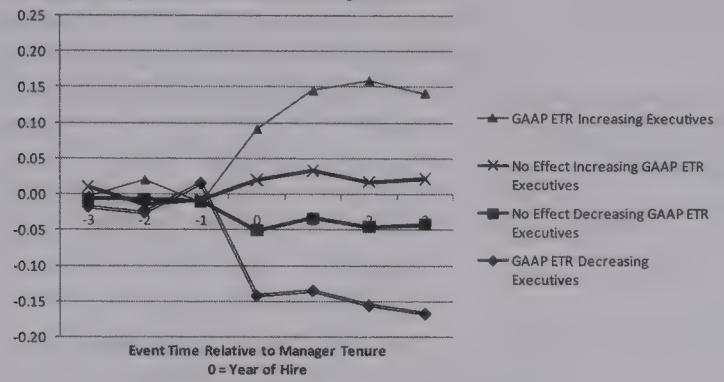
Figure 2, Panel B shows the corresponding results for the date of departure. Once the executives leave their firms, the *GAAP ETR* reverts back to levels they had been prior to the executives' employment. Together, Panels A and B show that if *GAAP ETR* were changing due to some factor other than the new executives, it would have to perfectly coincide with both the executives' hiring dates and the executives' departure dates to explain the data.

Panels C and D of Figure 2 present the changes in *CASH ETR* surrounding the dates of hire and departure, where the year of departure is $t = 0$. The pattern is similar to what we find in the *GAAP ETR* data. Once the executives join their firms, the *CASH ETR* begins to diverge. After the executives leave their firms, the *CASH ETR* reverts back to the levels that existed prior to the

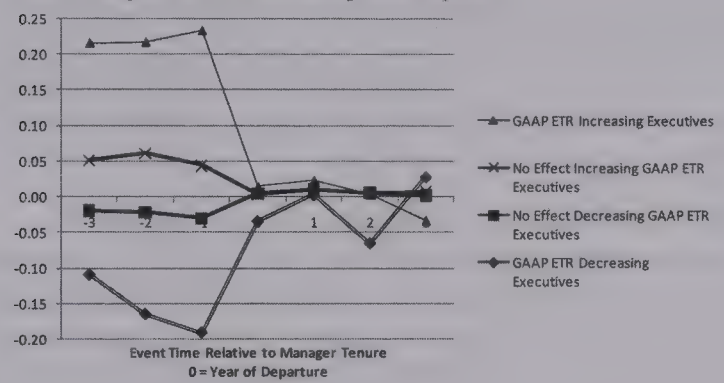
[15] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

FIGURE 2
Illustration of the Change in *GAAP ETR* and *CASH ETR* Surrounding the Year of Executive Hire and Departure

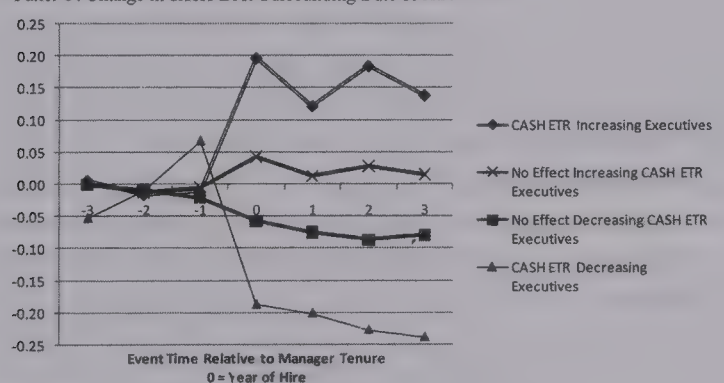
Panel A: Change in *GAAP ETR* Surrounding Date of Hire



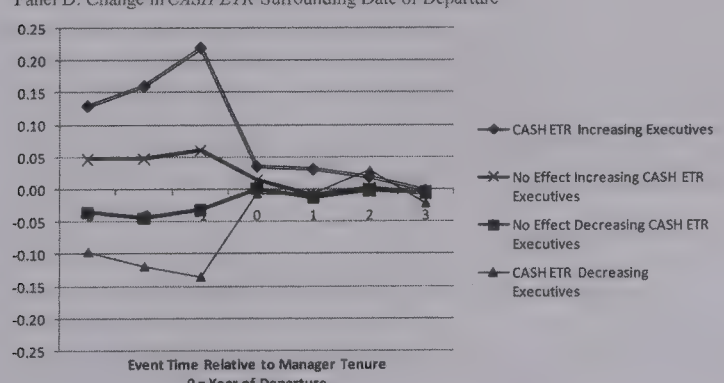
Panel B: Change in *GAAP ETR* Surrounding Date of Departure



Panel C: Change in *CASH ETR* Surrounding Date of Hire



Panel D: Change in *CASH ETR* Surrounding Date of Departure



The figure shows change in effective tax rates around the hiring (Panels A and C) and departure (Panels B and D) of the executive. Managers are assigned to decreasing, increasing, and no-effect groups by observing the manager's fixed effect coefficient. The no-effect groups are further split into those with positive insignificant coefficients and those with negative insignificant coefficients. Managers that have a statistically negative coefficient (at the 5 percent level) are labeled "decreasing" managers. Managers that have a statistically positive coefficient (at the 5 percent level) are labeled "increasing" managers. Managers whose coefficient is insignificant are labeled "no effect" managers. Each firm's *GAAP ETR* or *CASH ETR* is normalized using its average value of *GAAP ETR* or *CASH ETR* from the event-time period (−3 to −1).

executive joining the firm. Collectively, the results in Figure 2 illustrate the changes to *GAAP ETR* and *CASH ETR* when executives join firms and the return of *GAAP ETR* and *CASH ETR* to their prior levels once executives depart.

Our tests up to this point have focused on determining whether executives have a statistically significant incremental effect on their firms' reported effective tax rates, beyond firm fixed effects and time-varying firm characteristics. In the next section we examine whether the results are economically significant.

The Economic Magnitude of Executive Effects on Tax Avoidance

An appealing aspect of the empirical design (i.e., estimating a regression of an *ETR* on executive fixed effects) is that it generates a tax avoidance estimate for each executive in the sample. Thus, the coefficient on Executive A's fixed effect is an estimate of Executive A's influence on the effective tax rates of the firms where she/he has been employed.[16] In Table 4, we present the distribution of the executive fixed effect coefficients. In the table we include all executive fixed effect coefficients, be they CEO, CFO or other executives. The first two rows document the distribution of fixed effects for the primary variables of interest in the study, namely *GAAP ETR* and *CASH ETR*. The remaining rows show the distribution of estimated executive fixed effects for other economic outcomes. These additional effects are included in the table to show the economic magnitude of executive effects on other economic outcomes.

The results indicate that executives can have economically large effects on their firms' effective tax rates, in both directions. Among the entire set of executives, those at the 25th percentile showed a *GAAP ETR* reduction of 5.7 percent, while those at the 75th percentile were associated with a *GAAP ETR* increase of 5.4 percent. Since reductions in GAAP effective tax rates translate directly into increased reported earnings, executives effective at promoting tax avoidance have a significant effect on their firms' reported earnings that is proportionally larger than their impact on the GAAP effective tax rate.[17] For example, consider a firm that, given a certain level of pre-tax earnings and absent an executive specific effect, would have an effective tax rate of 40 percent. An executive who could reduce the effective tax rate by 5 percent would increase after-tax earnings by about 8.33 percent, as given by $(t - t') / (1 - t)$, where t is the effective tax rate with no executive effect and t' is the effective tax rate after the executive effect. The effects on *CASH ETR* are

[16] Naturally, these are simply estimates and are measured with error as indicated by the standard errors of the respective coefficients.

[17] We assume no implicit tax effect (cost to avoid the tax) in our computations but consider a given level of pre-tax earnings. We leave a detailed consideration of the consequences of executive determined tax avoidance to future research.

TABLE 4

Magnitude of Fixed Effects

Fixed Effect	n	Mean	Std. Dev.	P25	P50	P75
GAAP ETR	795	0.001	0.112	−0.057	−0.001	0.054
CASH ETR	783	−0.010	0.144	−0.080	−0.003	0.063
EBITDA	896	0.001	0.074	−0.034	0.003	0.032
R&D	891	−0.005	0.041	−0.009	−0.001	0.005
ADVERTISING	891	0.000	0.013	−0.003	−0.000	0.002
SG&A	660	−0.009	0.126	−0.036	−0.002	0.029
CAPITAL EXPENDITURES	751	0.005	0.194	−0.060	0.008	0.079
PERCENTAGE CHANGE IN SALES	908	0.003	0.253	−0.089	−0.002	0.098
LEVERAGE	891	0.004	0.089	−0.044	0.004	0.053
CASH HOLDINGS	891	0.003	0.123	−0.035	0.006	0.051
FOREIGN OPERATIONS	908	0.018	0.210	−0.052	0.002	0.070

The table presents the distribution of the executive fixed effects estimated in Table 3, Panels A and B (for *GAAP ETR* and *CASH ETR*, respectively). The estimation of the executive fixed effects for other variables was not reported. The first column represents the number of executives for which the effect was estimated. The remaining columns describe the distribution.

similar in magnitude to the effects on *GAAP ETR*. Executives at the 25th percentile led to a reduction in their firms' *CASH ETR* of 8.0 percent, while those at the 75th percentile led to an increase in their firms' *CASH ETR* of 6.3 percent. Thus, the results indicate that the effects that executives appear to have on *GAAP ETR* and *CASH ETR* can be economically significant.

The remaining rows in Table 4 present the distributions of executive effects on other outcomes. These results are generally in line with those in Bertrand and Schoar (2003). For example, results show that an executive at the 25th percentile of the distribution of executive effects on *EBITDA* reduces *EBITDA* by about 3.4 percent of assets, while an executive at the 75th percentile of the distribution of executive effects on *EBITDA* increases *EBITDA* by about 3.2 percent of assets. Below, we provide some initial evidence about whether executives who are effective at influencing tax avoidance have particular attributes and/or also tend to have predictable managerial styles. We turn to these issues in the next two sections.

Management Styles and Tax Avoidance

Prior research has found evidence that executives have different "styles." For example, some executives may emphasize growth strategies. In this section, we examine whether executives' tendencies toward tax avoidance are associated with other individual specific effects that the executives have on their firms' operational and financial decisions and outcomes. It is possible, for example, that executives who emphasize cost-cutting in general are the same executives who emphasize tax avoidance (i.e., as one way to cut costs). On the other hand, it is possible that the propensity for tax avoidance is independent of other decisions made in the firm and not associated with any other executive styles.

We begin by estimating Equation (1) with a variety of dependent variables other than *GAAP ETR* and *CASH ETR*. Following prior research and including a variable for foreign operations, we use the following operational and financial variables as dependent variables in separate Equation (1)-type regressions: *EBITDA*, *R&D*, *ADVERTISING*, *SG&A*, *CAPITAL EXPENDITURES*, *PERCENTAGE CHANGE IN SALES*, *LEVERAGE*, *CASH HOLDINGS*, and *FOREIGN OPERATIONS*

(defined in Table 2). Each of these dependent variables is selected because it is plausible *ex ante* that it could represent the type of activity in which a manager interested in tax avoidance would engage. For example, executives who minimize SG&A might be the type who cut all costs, including taxes. Further, executives with high intangibles (R&D spending) and/or more foreign operations may have more opportunities to reduce effective tax rates.

The analysis works as follows. We regress each variable, for example *CAPITAL EXPENDITURES*, on a vector of time-varying controls, firm, year, and executive fixed effects.[18] We then test whether the executive fixed effect coefficients that we estimate from these regressions are associated with the executive fixed effect coefficients from the tax regressions discussed earlier. For each pair of coefficients, we estimate a univariate regression of the following form:

$$\hat{\alpha}_m^Y = \beta_0 + \beta_1 \hat{\alpha}_m^Z + u_m, \tag{2}$$

where $\hat{\alpha}_m^Y$ is the estimated fixed effect coefficient for executive m estimated in Equation (1) when Y (which represents either *GAAP ETR* or *CASH ETR*) is the dependent variable in Equation (1), and $\hat{\alpha}_m^Z$ is the estimated fixed effect coefficient for executive m in Equation (1) when Z (which represents the nontax variables listed above) is the dependent variable in Equation (1).

It is important to note that these are inherently low-power tests, as the dependent variables are themselves coefficients from the earlier executive effects regressions and, thus, are noisy estimates. The noise, in a univariate regression, will lead to downward bias in an OLS estimation of β. We attempt to mitigate this problem by using a GLS estimation technique—weighting the observation by the inverse of the standard error of the independent variable (Bertrand and Schoar 2003).

Table 5 presents the results. Note that each entry in the table represents the coefficient from a *separate* univariate weighted least-squares regression. In general, the data are consistent with executive effects on *GAAP ETR* being positively associated with executive effects on accounting income, as evidenced by the positive coefficient on *EBITDA*. This suggests the possibility of an implicit tax effect whereby activities that are tax-favored also have low pretax rates of return (Scholes et al. 2009, Jennings et al. 2009), or it could be that executives who are not good at increasing operating income make up for their deficiency by being good at tax avoidance.

We find very little evidence of executive fixed effects on effective tax rates being correlated with executive fixed effects on other variables.[19] While this may seem surprising at first, the results are consistent with prior research that has documented a high degree of variability in effective tax rates across firms, even within the same industry (Dyreng et al. 2008).

Executive Characteristics and Tax Avoidance

Prior research has found some evidence that executives with different backgrounds have predictable differences in managerial style (e.g., Bertrand and Schoar 2003). Accordingly, we attempt to obtain biographical information about each of the executives (e.g., age, educational background, gender) in our sample. Frequently these data are missing from ExecuComp. For each executive with missing information, we hand-collect biographical information through Internet

[18] We do not report the regression results from each of these separate regressions in the interest of space. However, we note that, as one might expect, the effect of an executive on such items as *R&D* spending and leverage is greater than the executives' effect on the effective tax rate measures. For example, F-statistics on the executive fixed effects in the *R&D* regressions range (over executive ranks) from 1.0–5.0 and for leverage the range is 2.35–3.74. In comparison, the range for the tax regressions (from Table 3) is 1.36–2.15.

[19] In untabulated robustness tests, for each test in this table we exclude the one specific time varying control (e.g., *R&D*) from the right-hand side of Equation (1), where we have *GAAP ETR* or *CASH ETR* as the dependent variable, before we use that executive fixed effect in the Equation (2) tests for that specific variable (e.g., *R&D*). These robustness tests have little effect on the results. Across the 18 test statistics in the table, only three are significant at the 5 percent level.

TABLE 5

Correlation of *GAAP ETR* and *CASH ETR* Fixed Effects with Other Fixed Effects

Independent Variable	Dependent Variable	
Estimated Fixed Effect from Equation (1) for Manager *m* When Z is the Dependent Variable and Z is:	**Estimated Fixed Effect from Equation (1) for Manager *m* When *GAAP ETR* is the Dependent Variable**	**Estimated Fixed Effect from Equation (1) for Manager *m* When *CASH ETR* is the Dependent Variable**
EBITDA	0.321***	−0.028
	(0.056)	(0.072)
R&D	−0.158	−0.170
	(0.110)	(0.148)
ADVERTISING	0.633	0.504
	(0.355)	(0.460)
SG&A	−0.015	0.047
	(0.043)	(0.058)
CAPITAL EXPENDITURES	0.048**	0.048
	(0.024)	(0.030)
PERCENTAGE CHANGE IN SALES	0.029	−0.035
	(0.017)	(0.022)
LEVERAGE	−0.047	0.038
	(0.047)	(0.059)
CASH HOLDINGS	−0.017	0.063
	(0.034)	(0.045)
FOREIGN OPERATIONS	0.009	−0.043*
	(0.018)	(0.023)

*, **, *** Represent statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

This table presents tests of association between managers' "style" and their proclivity to avoid taxes. Each entry in the table represents the coefficient β_1 from the following model run separately (univariately) for each independent variable listed above: $\hat{a}_m^Y = \beta_0 + \beta_1 \hat{a}_m^Z + u_m$. In this model, $\hat{a}_m^Y$ is the estimated fixed effect for manager m when: Y (*GAAP ETR* or *CASH ETR*) is the dependent variable in Equation (1); $\hat{a}_m^Z$ is the estimated fixed effect for manager m when Z is the dependent variable in Equation (1); and Z is one of the following: *EBITDA*, *R&D*, *ADVERTISING*, *SG&A*, *CAPITAL EXPENDITURES*, *PERCENTAGE CHANGE IN SALES*, *LEVERAGE*, *CASH HOLDINGS*, *FOREIGN OPERATIONS* (all of which are defined in Table 2). The estimate of β_1 for each regression is presented in the table, with the respective standard error below in parentheses. Descriptive statistics for the dependent and independent variables are presented in Table 4.

searches for those executives. Often, we could obtain additional information about the executives using this approach. In addition, we gather information about the executives' wealth sensitivity to changes in the stock price of their firm, proxies for their degree of optimism and/or overconfidence, and whether they move between firms as part of a larger executive team.

Table 6 presents descriptive statistics on the biographical information for the executives in the sample having available data. Panel A presents the data for the subset of executives who have coefficients estimated when *GAAP ETR* is the dependent variable in Equation (1). Panel B presents the data for the subset of executives who have coefficients estimated when *CASH ETR* is the dependent variable in Equation (1). We find that about 20 percent of the sample has an M.B.A. degree from a top 10 institution (as defined by *BusinessWeek*).[20] Another 18 percent of executives

[20] Rankings can be found online at: http://bwnt.businessweek.com/interactive_reports/mba_domestic_2008/index.asp.

TABLE 6

Descriptive Statistics of Executive Characteristics

Panel A: Descriptive Statistics for Regression of Executive Fixed *GAAP ETR* Effects on Executive Characteristics

Variable	n	Mean	Std. Dev.	P25	P50	P75
TOP10 MBA	521	0.200	0.400	0.000	0.000	0.000
OTHER MBA	521	0.182	0.387	0.000	0.000	0.000
JURIS DOCTORATE	521	0.119	0.324	0.000	0.000	0.000
ACCOUNTING DEGREE	521	0.106	0.308	0.000	0.000	0.000
NO DEGREE RECORDED	521	0.075	0.263	0.000	0.000	0.000
CFO AT PRIOR FIRM	521	0.225	0.418	0.000	0.000	0.000
FEMALE	521	0.044	0.206	0.000	0.000	0.000
AVERAGE TENURE	521	5.036	1.463	4.000	5.000	6.000
AVERAGE AGE	521	50.880	6.761	46.000	50.769	55.359
SAME INDUSTRY SWITCH	521	0.438	0.497	0.000	0.000	1.000
MOVES AS A TEAM	521	0.148	0.355	0.000	0.000	0.000
OVERCONFIDENCE IN FORECASTS	521	0.399	0.269	0.200	0.400	0.600
LOG OF WEALTH SENSITIVITY TO PRICE	521	9.283	1.917	8.099	9.291	10.644
OPTIMISM IN 10K	521	0.000	0.031	−0.018	0.000	0.016

Panel B: Descriptive Statistics for Regression of Executive Fixed *CASH ETR* Effects on Executive Characteristics

Variable	n	Mean	Std. Dev.	P25	P50	P75
TOP10 MBA	513	0.203	0.402	0.000	0.000	0.000
OTHER MBA	513	0.181	0.386	0.000	0.000	0.000
JURIS DOCTORATE	513	0.119	0.324	0.000	0.000	0.000
ACCOUNTING DEGREE	513	0.107	0.310	0.000	0.000	0.000
NO DEGREE RECORDED	513	0.076	0.265	0.000	0.000	0.000
CFO AT PRIOR FIRM	513	0.224	0.417	0.000	0.000	0.000
FEMALE	513	0.045	0.207	0.000	0.000	0.000
AVERAGE TENURE	513	5.043	1.464	4.000	5.000	6.000
AVERAGE AGE	513	50.904	6.785	46.000	50.833	55.442
SAME INDUSTRY SWITCH	513	0.435	0.496	0.000	0.000	1.000
MOVES AS A TEAM	513	0.150	0.358	0.000	0.000	0.000
OVERCONFIDENCE IN FORECASTS	513	0.399	0.270	0.200	0.403	0.600
LOG OF WEALTH SENSITIVITY TO PRICE	513	9.287	1.925	8.099	9.294	10.647
OPTIMISM IN 10K	513	0.000	0.031	−0.018	0.000	0.016

have an M.B.A. from other than top 10 institutions. About 12 percent of the executives have law degrees, 11 percent have accounting degrees, and about 7 percent have no degree recorded. We focus on M.B.A., J.D., and accounting degrees because it is possible that executives with financial or legal training might be more likely to emphasize tax avoidance. About 23 percent of the sample was CFO at his/her prior firm (unconditional on current title), and about 5 percent of the sample is female. The average executive's tenure is about five years per firm. Executives average about 51 years of age during their tenure. About 44 percent of our sample executives switch from one firm to another within the same two-digit SIC code, and about 15 percent eventually end up working with the same executives at both firms (although the years do not overlap perfectly so as to allow for statistical estimation in the first-stage model).

We also include estimates of the executive's overconfidence, wealth sensitivity, and optimism. We measure executive overconfidence as the average frequency the firm missed its own managerial forecast during the years the executive was employed (Hribar and Yang 2007). We calculate the executive's wealth sensitivity to a percentage change in stock price following Core and Guay (1999). To measure optimism, we download all 10-K filings for all firm-years in our sample from http://www.sec.gov. We then use the software "General Inquirer" to count the number of positive and negative words in the 10-K (see Kothari et al. [2009] for a recent study that uses the General Inquirer software in an accounting disclosure context). We use the ratio of positive words to negative words as a measure of optimism. We subtract each firm-year value of optimism from the firm's average optimism value. For each executive, we take the average of the optimism scores for the firm-years for which s/he has sufficient data in our sample. Thus, the executive's optimism score is his/her average optimism at the firms at which she/he was employed during the years she/he was employed, adjusting for each firm's mean optimism.

Table 7 presents the results of regressions of executive fixed effect coefficients obtained from the empirical estimation of Equation (1) on variables that capture the biographical characteristics from Table 6. The first two columns present the results using executive fixed effect coefficients from the *GAAP ETR* model of Equation (1) as the dependent variable. The second two columns present the results using the executive fixed effect coefficients from the *CASH ETR* model of Equation (1) as the dependent variable.

Executive backgrounds do not explain much of the variation in tax avoidance across executives. We find no relation between the educational background of the executives in our sample and their propensities to reduce effective tax rates. The lack of connection with financial educational background suggests that executives' effects on tax avoidance is likely due more to the "tone at the top" than it is to direct involvement by the executives in the tax avoidance.

We also test whether the executive effects depend on whether the executive moved to a firm in a different industry or stayed in the same industry, and whether the executive moved alone or as part of a team. There is weak evidence that staying in the same industry is associated with less executive-specific tax avoidance in the sense of higher *CASH ETR*, but there is no significant difference for *GAAP ETR*. The results indicate no difference in the executive effects between moving alone and moving as a team. We find weak evidence of a link between optimism in the firm's 10-K and lower cash effective tax rates. That is, it appears that more optimistic executives reduce *CASH ETR*s to a greater extent. There is no evidence that the executive's level of overconfidence influences his/her ability to affect effective tax rates. Finally, we examine the executive's compensation structure and find no significant association with the executive fixed effects.

Overall, the results indicate that biographical information such as educational background and age does not explain much of the variation in tax avoidance across executives. We note that this finding does not suggest that executives have no effect on tax avoidance. Instead, these results are evidence that common, observable characteristics are not strongly associated with executives' propensities to reduce effective tax rates.

V. CONCLUSIONS

We track the movement of 908 CEOs, CFOs, and other executives across firms during the period 1992 to 2006 in order to examine whether individual executives have an effect on their firms' tax avoidance. By examining executives who switch firms, we attempt to control for firm fixed effects and identify executive-specific effects. Results indicate that individual executives play a significant role in determining the level of tax avoidance that firms undertake, incremental to characteristics of the firm. Analyzing executive effects in event time shows results consistent with executive-specific tax avoidance that arises shortly after the executive arrives and stops after the executive leaves. In addition, an executive's effect on tax avoidance at his/her first firm tends

TABLE 7

Tax Fixed Effects on Manager Characteristics

	GAAP ETR Fixed Effects		*CASH ETR* Fixed Effects	
Parameter	**Estimate**	**t-stat**	**Estimate**	**t-stat**
Intercept	−0.022	−0.53	−0.044	−0.99
TOP10 MBA	−0.014	−1.10	0.005	0.30
OTHER MBA	0.000	0.00	−0.003	−0.22
JURIS DOCTORATE	0.007	0.45	0.012	0.69
ACCOUNTING DEGREE	−0.007	−0.45	0.011	0.43
NO DEGREE RECORDED	−0.009	−0.48	−0.006	−0.39
CFO AT PRIOR FIRM	−0.013	−1.12	0.016	1.05
FEMALE	0.009	0.39	0.035	1.34
AVERAGE TENURE	−0.002	−0.64	0.003	0.93
AVERAGE AGE	0.000	0.43	0.001	1.33
SAME INDUSTRY SWITCH	0.007	0.72	0.021*	1.89
MOVES AS TEAM	−0.009	−0.63	0.021	1.23
OVERCONFIDENCE IN FORECASTS	0.004	0.25	−0.012	−0.54
LOG OF WEALTH SENSITIVITY	0.003	0.96	−0.004	−1.44
OPTIMISM IN 10K	−0.107	−0.71	−0.402**	−2.14
n	521		513	
R^2	0.012		0.025	
ADJRSQ	−0.015		−0.002	

*, **, *** Represent statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

This table presents results from testing the association of manager characteristics with manager fixed effects from Equation (1). The model is:

$$\hat{\alpha}_m^Y = \omega_0 + \omega_k CHARACTERISTIC_m^k + \varepsilon_m,$$

where: $\hat{\alpha}_m^Y$ is the estimated fixed effect for manager m when Y (*GAAP ETR* in Panel A or *CASH ETR* in Panel B) is the dependent variable in Equation (1); and $CHARACTERISTIC_m^k$ is the kth characteristic of manager m. We include the following characteristics: *TOP10 MBA* indicates the manager earned his M.B.A. from a business school ranked in the top 10 in the 2008 issue of *BusinessWeek* (available at: http://www.businessweek.com/bschools/rankings/).

Variable Definitions:

OTHER MBA = the manager earned his M.B.A. from a business school other than one of the top 10 business schools;
JURIS DOCTORATE = the manager earned a Juris Doctor;
ACCOUNTING DEGREE = at least one of the manager's degrees was in accounting or had an accounting emphasis;
NO DEGREE RECORDED = no information was available about the manager's education;
CFO AT PRIOR FIRM = the manager was a CFO at his prior place of employment;
AVERAGE TENURE = the average number of years the manager spent at each company recorded in ExecuComp;
AVERAGE AGE = the average age of the manager during his tenure as an executive;
SAME INDUSTRY SWITCH = the manager switched firms within the same two-digit SIC code;
OPTIMISM IN 10K = the average level of optimistic words to pessimistic words in the annual 10-K reports for the firm-years during which the manager worked at the firm (adjusted for the firm mean). We used the software General Inquirer to classify words as optimistic or pessimistic. Documentation is available at: http://www.wjh.harvard.edu/~inquirer/;
OVERCONFIDENCE IN FORECASTS = the number of times the firm missed its own earnings forecast relative to the number of forecasts issued; management earnings forecasts were gathered from First Call; and
LOG OF WEALTH SENSITIVITY = the log of the manager's wealth sensitivity to a 1 percent change in stock price, calculated as in Core and Guay (1999) and averaged over the executive's tenure at both firms; we use iteratively reweighted robust regression to control for extreme observations.

to be positively associated with his/her effect at the second firm, mitigating potential concerns that our results are driven by random effective tax rate changes at only one firm during the executive's tenure.

Overall, our study documents that individual executives exhibit different proclivities toward tax avoidance. This variation in tax avoidance proclivity across executives affects their firms' tax avoidance in ways that cannot be explained by firm characteristics. Moreover, the evidence indicates that these executive effects can be economically large. Given the difficulty of explaining effective tax rates or other measures of tax avoidance, documenting that executives matter to their firms' effective tax rates is an important first step and contribution.

These findings raise many questions that we hope will be addressed by future research. It would be interesting to know what happens to these executives and the firms they manage in the future. Perhaps the executives who appear to emphasize tax avoidance simply lead their firms to take on substantial tax risk. In that case, we might expect to see additional taxes being paid in future years as the firms undergo auditing by the IRS, perhaps on another executive's watch. What happens to the career prospects of executives who emphasize or de-emphasize tax avoidance? In addition, future research could examine whether tax avoidance indicates anything about the executives' propensity to emphasize other behavior, including behavior that may not be in the best interests of the firm. These are just a few of the questions raised by this study that we believe can lead to further insights.

REFERENCES

Armstrong, C., J. Blouin, and D. Larcker. 2009. The incentives for tax planning. Working paper, University of Pennsylvania and Stanford University.

Bamber, L., J. Jiang, and I. Wang. 2010. What's my style? The influence of top managers and their personal backgrounds on voluntary corporate financial disclosure. *The Accounting Review* 85 (4): 1131–1162.

Bennedsen, M., F. Perez-Gonzalez, and D. Wolfson. 2008. Do CEOs matter? Working paper, New York University.

Bertrand, M., and A. Schoar. 2003. Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics* 68 (4): 1169–1208.

Callihan, D. 1994. Corporate effective tax rates: A synthesis of the literature. *Journal of Accounting Literature* (13): 1–43.

CFO Research Services and Deloitte Tax LLP. 2006. *What do Companies Want from the Corporate Tax Function? CFO and Tax Executives' Perspectives on Corporate Tax.* Boston, MA: CFO Publishing Corp.

Chatterjee, A., and D. Hambrick. 2007. It's all about me: Narcissistic CEOs and their effects on company strategy and performance. *Administrative Science Quarterly* 52 (3): 351–386.

Chen, S., X. Chen, Q. Cheng, and T. Shevlin. 2010. Are family firms more tax aggressive than non-family firms? *Journal of Financial Economics* 95 (1): 41–61.

Core, J., and W. Guay. 1999. The use of equity grants to manage optimal equity incentive levels. *Journal of Accounting and Economics* 28 (2): 151–184.

Crocker, K., and J. Slemrod. 2005. Corporate tax evasion with agency costs. *Journal of Public Economics* 89: 1593–1610.

Desai, M., and D. Dharmapala. 2008. Corporate tax avoidance and firm value. *The Review of Economics and Statistics.*

———, and ———. 2006. Corporate tax avoidance and high-powered incentives. *Journal of Financial Economics* 79 (1): 145–179.

Drucker, J. 2007. Inside Wal-Mart's bid to slash state taxes; Ernst & Young devises complex strategies; California pushes back. *Wall Street Journal* (October 23): A1.

Dyreng, S., M. Hanlon, and E. Maydew. 2008. Long run corporate tax avoidance. *The Accounting Review* 83 (1): 61–82.

Frank, M., L. Lynch, and S. Rego. 2009. Tax reporting aggressiveness and its relation to aggressive financial reporting. *The Accounting Review* 84 (2): 467–498.

Ge, W., D. Matsumoto, and J. Zhang. 2009. Do CFOs have styles of their own? An empirical investigation of the effect of individual CFOs on financial reporting practices. Working paper, University of Washington.

Gompers, P., J. Ishii, and A. Metrick. 2003. Corporate governance and equity prices. *The Quarterly Journal of Economics* 118: 107–155.

Graham, J. 2003. Taxes and corporate finance: A review. *Review of Financial Studies* (16): 1074–1128.

———, and A. Tucker. 2006. Tax shelters and corporate debt policy. *Journal of Financial Economics* 81: 563–594.

Gupta, S., and K. Newberry. 1997. Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data. *Journal of Accounting and Public Policy* 16 (1): 1–34.

Hanlon, M., and S. Heitzman. 2010. A review of tax research. *Journal of Accounting and Economics* (forthcoming).

Hanlon, M., and J. Slemrod. 2009. What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. *Journal of Public Economics* 93 (January): 126–141.

Hayes, R., and S. Schaefer. 1999. How much are differences in managerial ability worth? *Journal of Accounting and Economics* 27 (2): 125–148.

Hribar, P., and H. Yang. 2007. CEO confidence, management earnings forecasts, and earnings management. Working paper, The University of Iowa and Cornell University.

Jennings, R., W. Mayew, and C. Weaver. 2009. The effect of TRA86 on the extent of implicit taxes at the corporate level. Working paper, Duke University, The University of Texas at Austin, and Texas A&M University.

Johnson, B., R. Magee, N. Nagarajan, and H. Newman. 1985. An analysis of the stock price reaction to sudden executive death: Implications for the management labor market. *Journal of Accounting and Economics* 7: 151–174.

Kothari, S., X. Li, and J. Short. 2009. The effect of disclosures by management, analysts, and financial press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. *The Accounting Review* 84 (5): 1639–1670.

Lisowsky, P. 2009. Seeking shelter: Empirically modeling tax shelters using financial statement information. *The Accounting Review* (forthcoming).

Liu, C., and D. Yermack. 2008. Where are the shareholders' mansions? CEOs home purchases, stock sales, and subsequent company performance. Working paper, Arizona State University and New York University.

Malmendier, U., and G. Tate. 2009. Superstar CEOs. *The Quarterly Journal of Economics* 124 (4): 1593–1638.

Mills, L., M. Erickson, and E. Maydew. 1998. Investments in tax planning. *The Journal of the American Taxation Association* 20 (1): 1–20.

Murphy, K., and J. Zimmerman. 1993. Financial performance surrounding CEO turnover. *Journal of Accounting and Economics* 16 (1–3): 273–316.

Petersen, M. 2009. Estimated standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.

Phillips, J. 2003. Corporate tax planning effectiveness: The role of compensation-based incentives. *The Accounting Review* 78 (3): 847–874.

Porcano, T. 1986. Corporate tax rates: Progressive, proportional, or regressive. *The Journal of the American Taxation Association* 7: 17–31.

Pourciau, S. 1993. Earnings management and nonroutine executive changes. *Journal of Accounting and Economics* 16 (1–3): 317–336.

Rego, S. 2003. Tax-avoidance activities of U.S. multinational corporations. *Contemporary Accounting Research* 20 (4): 805–833.

———, and R. Wilson. 2009. Executive compensation, tax reporting aggressiveness, and future firm performance. Working paper, The University of Iowa.

Robinson, J., S. Sikes, and C. Weaver. 2010. The impact of evaluating the tax function as a profit center on

effective tax rates. *The Accounting Review* 85 (3): 1035–1064.

Scholes, M., M. Wolfson, M. Erickson, E. Maydew, and T. Shevlin. 2009. *Taxes and Business Strategy: A Planning Approach*. 07458 Upper Saddle River, NJ: Pearson Prentice Hall.

Shackelford, D., and T. Shevlin. 2001. Empirical tax research in accounting. *Journal of Accounting and Economics* (31): 321–387.

Shevlin, T., and S. Porter. 1992. The corporate tax comeback in 1987: Some further evidence. *The Journal of the American Taxation Association* 14: 58.

Siegfried, J. 1974. Effective average U.S. corporation income tax rates. *National Tax Journal* (27): 245–259.

Stickney, C., and V. McGee. 1982. Effective corporate tax rates, the effect of size, capital intensity, leverage and other factors. *Journal of Accounting and Public Policy* (1): 23–45.

Wilson, R. 2009. An examination of corporate tax shelter participants. *The Accounting Review* 84 (3): 969–999.

Zimmerman, J. L. 1983. Taxes and firm size. *Journal of Accounting and Economics* (5): 119–149.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1191–1214

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1191

Oligopoly, Disclosure, and Earnings Management

Mark Bagnoli
Susan G. Watts
Purdue University

ABSTRACT: We examine how biased financial reports (managed earnings) affect product market competition and how product market competition affects incentives to bias financial reports in a model with fully rational firms. We find that Cournot competitors bias their reports to create the impression that their costs are lower than they actually are. This bias leads to lower total production and a higher product price, even though each firm fully understands its rival's incentives to bias its financial reports. The magnitude of the bias is larger when firms compete in more profitable product markets and smaller when the firm can extract more information about its rival's costs from its own. When the costs of misreporting are asymmetric, the lower-cost firm engages in more earnings management than its rival, produces more than it would in a full-information environment, and earns greater profits. Our analysis leads to new, testable relations among earnings management, reported and actual earnings, and industry structure.

Keywords: *earnings management; reporting bias; product market competition; information asymmetry.*

JEL Classifications: *G1; G14; M41.*

I. INTRODUCTION

We develop a theoretical model to examine how earnings management affects product market competition and how competition affects earnings management decisions. A key feature of our analysis is that, in addition to an information extraction problem, there is an important component of rivalry in the firms' earnings management and production decisions. The motivation for our study arises, in part, from claims made by C. Michael Armstrong who was CEO of AT&T from 1997 to 2002, one of the most turbulent times for the company. He has recently argued that accounting fraud at WorldCom was the cause of AT&T's *perceived*

We thank Steve Kachelmeier (senior editor), Harry Evans (editor) and an anonymous referee for suggestions and guidance that have significantly improved this study. We also thank Jennifer Altamuro, Matt Anderson, Anil Arya, Ramji Balakrishnan, Jeff Burks, Ricky Cazier, Leslie Hodder, Pat Hopkins, Xu Jiang, Mike Kirschenheiter, Kay Li, Scott Liao, Dara Marshall, Eric Rasmussen, Ram Venkataraman, and participants at the 2007 Midwest Accounting Research Conference for helpful comments. Finally, we thank Hsin-Tsai Liu for invaluable research assistance and gratefully acknowledge the financial support provided by the Krannert Graduate School of Management and Purdue University.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: May 2008
Accepted: October 2009
Published Online: June 2010

strategic failures, its inability to compete with WorldCom and, in the end, the decision to break up the company. He specifically suggests that WorldCom's fraudulently reported revenues, margins and costs drove AT&T's layoffs, cost-cutting, and a very unprofitable price war that left AT&T unable to service the debt he incurred to revive the company. This view is supported by William Estry who was Sprint's CEO during that same time.[1]

To address our research questions, we employ an incomplete information Cournot duopoly model in which each firm knows its own production costs, but not its rival's. In our model, each firm provides a disclosure through, for example, its income statement. A firm's rival can use the disclosure to update its beliefs about the disclosing firm's production costs prior to competing in the product market. Our model differs from prior work on disclosure in incomplete information Cournot models because we assume that firms can provide biased reports and that all firms optimally update their beliefs using all available information.[2] If firms bias their reports, they incur a cost of misreporting. We show that when the firms know each other's cost of misreporting, there are no linear equilibria. However, when these costs are also private information, there are linear equilibria in which the firms bias their disclosures to gain a competitive advantage in the product market. Thus, our study complements and extends the results in Fischer and Verrecchia (2004). In their model, only one firm has private information, and biased reporting in equilibrium requires that some firms use heuristics when adjusting their beliefs (i.e., some firms are not Bayesian). Our model differs from theirs in two important ways. First, we allow for rivalry to impact earnings management decisions (that is, more than one firm has private information) and second, because all firms optimally update their beliefs (i.e., all firms are Bayesian), we extend the literature by showing that biased reporting can occur in equilibrium in a model with fully rational firms.

Intuitively, in our model, each firm's disclosure is designed to create the impression that its production costs are lower than they actually are. Further, because each firm uses all available information efficiently and fully understands its rival's incentives to bias its report, each adjusts its beliefs about its rival's production costs upward relative to the disclosure, but still underestimates those costs in equilibrium. Consequently, each firm cuts its production below the full-information level, causing the market price to rise and each firm's product market profits to increase, in turn providing just the incentive needed for each firm to bias its disclosure.[3] Our analysis indicates that these effects are smaller when firms compete in more profitable product markets or when they use reasonably similar technologies. Interestingly, while the effects are smaller in both cases, they

---

[1] See, for example, the book written in 2004 by Dick Martin, the former head of public relations at AT&T (Martin 2004), or the articles by Searcy (2005), Blumenstein and Grant (2004), and McConnell et al. (2002). While many are familiar with the most famous aspects of WorldCom's accounting fraud (i.e., capitalizing line costs and booking non-existent revenues), there appears to be no bright line between earnings management and fraud (Mulford and Comiskey 2002, 41; Ronen and Yaari, 2008, 25–31; Schilit 1993, 1; or the survey by Beneish 2001). This is probably due, in part, to the inherent difficulties in adequately defining earnings management and distinguishing it from fraud. In WorldCom's case, its manipulation of earnings began slowly and became more egregious, culminating in actions that virtually all agree are fraud. Some of the early activities that may or may not be fraudulent include recording and releasing reserves and capitalizing internal labor costs related to line installation (Thornburgh 2002).

[2] The prior work using incomplete information Cournot models includes Fried (1984), Shapiro (1986), Gal-Or (1986, 1988), Darrough (1993), Raith (1996), and Vives (2002). Also see the surveys by Verrecchia (2001) and Vives (2006). When these authors focus on disclosure, they assume that the firms' disclosures are unbiased but that the firms can determine the amount of noise in the disclosure. Thus, if the firm wishes to disclose, then it chooses not to add noise and if it wishes not to disclose, it chooses to add an infinite amount of noise.

[3] As a result, the transmission mechanism that produces value to biasing reports is different in our setting than in Fischer and Verrecchia (2004). In their setting, the Bayesian firms recognize that their rivals employ heuristics and thus "overproduce" relative to the full-information setting. Realizing this, the Bayesian firms accommodate the increased production by reducing theirs. In our model, a firm's biased report leaves its rivals rationally believing that the reporting firm's costs are lower than they actually are, which leads the rival to believe that the reporting firm will produce more than it actually does. As a result, in a symmetric equilibrium, total output is less than the full-information level and so *both* firms benefit.

differ in that firms engage in more (less) misreporting in the first (second) situation. Finally, when firms have sufficiently different costs of misreporting, the firm with lower costs introduces greater bias into its reports, produces more output than it would in the full-information outcome, and earns greater product market profits. Thus, our analysis supports Armstrong's view of the competitive impact of WorldCom's fraudulent accounting and suggests that it may be worth examining similar competitive environments for similar activity.

By focusing on equilibrium incentives to bias reports, our analysis suggests new empirical implications associated with the use of different earnings management techniques. First, only some such techniques effectively bias inferences about the firm's production costs.[4] Thus, our model predicts that the use of these techniques is positively correlated with standard measures of profitability such as ROE calculated using original (unadjusted) financial statement information but negatively correlated with these measures calculated using restated (adjusted) financial statement information, if available.[5] In addition, if incentives to use other earnings management techniques are orthogonal to the product market incentives that are the focus of our analysis, then increases in the cost of these latter techniques (e.g., as the result of the passage of the Sarbanes-Oxley Act of 2002) should reduce the observed proportions of their use among SEC enforcement actions. Third, we show that firms with similar production technologies will be less likely to employ these types of earnings management techniques. Thus, firms whose production is governed by physical or chemical processes or firms in mature industries are less likely to engage in this type of earnings management than firms in service industries or firms that have significantly larger portfolios of products. Finally, our analysis suggests that the increased use of deferred prosecution agreements (DPAs) is likely to lead to unexpected consequences. In particular, because such agreements dramatically increase the cost of misreporting, competitors of firms operating under DPAs are likely to respond by increasing the amount of misreporting in order to gain/increase their competitive advantage in the product market.

In addition to contributing to the literature on the interaction between disclosure and product market competition, our study also highlights and extends a key insight of Fischer and Verrecchia (2000) into biased disclosures. They introduce uncertainty about the manager's objective function into a standard voluntary disclosure model (e.g., Verrecchia 1983) and show that the manager optimally introduces bias into the disclosure. Essentially, the agency problem combines with incomplete contracting to limit the market's ability to fully adjust its expectation of firm value leaving the manager with a personal benefit to offering a biased voluntary disclosure. In our model, the additional uncertainty associated with the market being unsure of the firm's reporting costs opens a similar opportunity for the firm to bias its disclosure, but does not rely on the combination of an agency problem and incomplete contracting. Thus, our model enriches the understanding of biased disclosures by showing that they may arise even if all agency problems and incomplete contracting issues are eliminated.

The literature on the interaction between disclosures and product market competition is extended and synthesized in Darrough (1993), generalized in Raith (1996), and summarized in

---

[4] Examples of this type of earnings management include fraudulent or aggressive revenue recognition, aggressive cost capitalization, inclusion of operating costs in restructuring costs, inappropriately low estimates of the allowance for doubtful accounts or warranty expense, and certain types of "real" earnings management such as delaying expenditures. Examples of earnings management techniques that would not be expected to lead to biased inferences about production costs include channel stuffing, delayed write-downs of assets, and the "timely" sale of assets.

[5] Since adjusted financial statement information is generally only available for firms that have restated earnings, this result is likely to be most easily tested by focusing on the subset of firms that have restated earnings and admitted to employing earnings management techniques that effectively bias inferences about production costs.

Vives (2006).[6] The key difference between this literature and our model is that prior work required firms to make unbiased disclosures—firms could only alter the amount of noise in their disclosure—whereas we permit firms to bias their disclosures at a cost. This difference leads to very different predictions. For Cournot competitors with private information about production costs, Darrough (1993) and others have shown that firms prefer to precommit to disclosing their private information without adding any noise. Intuitively, adding noise results in the firm's rival choosing to produce more than it would if fully informed, which lowers price and product market profits.[7] If the disclosure decision is made after firms learn their production costs, then because they are required to disclose truthfully, the standard unraveling story holds.[8] Our results complement this literature by showing that when fully rational firms are permitted to bias disclosures (even at a cost), they choose to do so and that this decision is affected by and affects competition in the product market.[9]

Our analysis is also related to recent work by Sadka (2006) and Kedia and Philippon (2009). Sadka (2006) examines accounting fraud in a competitive (price-taking) model in which fraud is detectable unless the firm alters its output decisions to mimic a low-cost producer. He shows that accounting fraud lowers social welfare because the fraudulent firms produce too much output at too high a cost and the low-cost firms produce too little relative to the Pareto Efficient outcome. Kedia and Philippon (2009) examine accounting fraud in a signaling framework in which the fraudulent firm must overinvest and overhire in order to mimic the more efficient producers. They exploit this structure to suggest why the detection of accounting fraud can impact not only the firm's rivals, but can have economy-wide effects, too. Thus, while the models differ, the central idea is that fraud can be detected by examining a firm's real (as opposed to financial) decisions and that this is the transmission mechanism that links reporting activity to real effects. Our analysis complements this approach by showing that there is an additional linkage by which firms can influence their competitive position in the product market by introducing bias into their financial reports.

Section II contains a description of our model of disclosure and product market competition. In Section III, we discuss equilibrium in restricted versions of our model. In Section IV, we focus on symmetric equilibria and analyze asymmetric equilibria in Section V. We conclude in Section VI.

II. THE MODEL AND EQUILIBRIUM

In this section, we develop a model designed to provide insight into how earnings management affects product market competition and how competition affects earnings management decisions. To address these questions, we need to fully capture the idea that each firm understands that its rival will try to extract an information advantage from the firm's financial reports while

---

[6] There is also an important literature on product market competition as a source of the cost of voluntary disclosures. This literature frequently focuses on the effects on a rival's entry decision from an incumbent firm's voluntary disclosure, thus endogenizing the cost of voluntary disclosure. The foundational studies in this literature include Darrough and Stoughton (1990), Wagenhofer (1990), and Feltham and Xie (1992).

[7] In a recent study, Sadka (2004) explores how disclosure of cost information, especially by the low-cost producer, can affect a firm's competitive advantage in a product market and exploits this model to show that the key results in Darrough (1993) are reinforced by the existence of his endogenized costs of disclosure.

[8] See Christensen and Feltham (2002) for a nice summary of this literature. Interestingly, Arya et al. (2010) show that if firms compete in multiple markets, then the unraveling result may fail. In particular, they show that there can be a partial pooling equilibrium in which firm types that have high costs in one market and low costs in the second pool with firm types with low costs in the first and high costs in the second.

[9] The issue of whether the firms precommit to disclose does not arise substantively in our model because firms can select a disclosure that is independent of their private information. Thus, even if firms precommit to disclose, they have an available strategy that allows them to reveal none of their private information.

simultaneously understanding that its rival realizes that the firm is trying to do the same thing—extract an information advantage from the rival's financial reports. Thus, a key feature is that not only is there an information extraction problem, but there is also an important component of rivalry in determining how each firm manages its earnings (in addition to the rivalry associated with the firms' competition in the product market).

To adequately capture these forces, both firms must be privately informed about their own costs and these costs must be correlated. Both firms having private information is required in order to analyze the competitive component of their earnings management decisions. If there is only one privately informed firm, then it is the only firm that can engage in earnings management designed to affect competition in the product market and so competition to manage earnings would be eliminated by assumption. The reason the firms' costs need to be correlated is somewhat more subtle, but can be most clearly seen by considering the impact of assuming the opposite—that their costs are not correlated. If this is case, then a firm's private information is not helpful in estimating its rival's costs, and thus the firm's financial reports *only* provide the rival with information about the firm's costs of production. Consequently, the firm's report can only impact competition in the product market—there is no competitive effect in the firms' earnings management decisions. However, when the firms' costs are correlated, the information one firm extracts about its rival's costs not only informs it about those costs, but also provides additional information about the rival's ability to extract information from the firm's own report.[10]

To address our research questions, we examine a two-stage game that captures the interactions between firm disclosures, earnings management decisions, and competitive pressures in the product market. In the first stage, each of two firms chooses a disclosure without knowing its rival's choice and incurs a disclosure cost. Each firm makes this disclosure knowing both its own disclosure and production costs. In the second stage, the firms observe both disclosures from the first stage and compete in the product market by choosing an amount of the homogeneous product to offer for sale without observing the output choice of its rival. We simplify the analysis by assuming that each firm has constant marginal costs of production and fixed costs are zero. Together, these imply that the firm's average costs of production equal its marginal costs.

In the disclosure stage, each firm chooses a disclosure that we interpret as a disclosure about its average cost of production. Our objective is to capture the idea that a firm's financial statements (e.g., an income statement) can be used to infer the firm's reported marginal/average costs. Further, since firms can make reporting decisions that either enhance or diminish reported earnings (both within the discretion afforded by GAAP or through fraud), we do not require truthful disclosure. Instead, we assume that a firm that does not disclose truthfully incurs a disclosure cost $h_i(s_i, c_i)$, $i = 1,2$, where s_i is firm i's disclosure and c_i is its cost of production. To keep the analysis tractable, we assume that $h_i(s_i, c_i) = k_i(s_i - c_i) + (1/2)\varepsilon_i(s_i - c_i)^2$, with $\varepsilon_i > 0$.[11] Notice that with this specification, firms that disclose truthfully incur no disclosure costs, but firms that do not incur two types of costs. The first type (represented by the second term) imposes a quadratic cost of misreporting c_i—if the firm misreports $(s_i \neq c_i)$, it incurs a cost that is proportional to the

[10] As an example, consider the case when each firms' costs of production are made up of two components, a common and an idiosyncratic component. If the firms' costs are not correlated, then the common component must be common knowledge and all each firm needs to do is extract its rival's idiosyncratic component from its rival's financial report. However, when the common component is not known with certainty, then the firm itself will be unsure of the split between the common and its idiosyncratic components and will use its *rival's* financial report to help it understand its own split. This will, in turn, impact its inference about its rival's production costs and the rival will need to factor this effect in when determining whether (or how much) to manage its earnings.

[11] Throughout, we assume that these costs are not prohibitively large.

squared difference between its disclosure and its actual marginal cost of production.[12] As a result, it is symmetric in deviations from disclosing truthfully. The second type of disclosure cost (represented by the first term) is not symmetric and depends on the sign of k_i. Intuitively, if $k_i > 0$, then the firm's disclosure costs increase when its disclosure exceeds its actual marginal cost of production and may represent an added disadvantage from the market inferring that it is less competitive. Similarly, if $k_i < 0$, then the firm's disclosure costs increase when its disclosure is smaller than its actual marginal costs and may represent an added disadvantage associated with a (perceived) weakened position in bargaining with employees or in dealing with regulators (Watts and Zimmerman 1986, 1990).

In the production stage of the game, the firms observe the first-stage disclosures (s_1, s_2), draw inferences from these disclosures optimally and compete in the product market. We model the product market using a fairly standard incomplete information, Cournot duopoly model where firms know their own costs of production but do not know their rival's.[13] More specifically, we assume that market demand for the homogeneous product sold by the two firms is $P = a - q_1 - q_2$ and units of output are normalized so that the slope coefficient is 1. All information about market demand is common knowledge.

An important component of our model is the information structure. We assume that firm i knows its own disclosure costs (k_i) and its costs of production (c_i) but not its rival's disclosure costs (k_j) or its rival's costs of production (c_j). All other parameters are common knowledge. Thus, each firm has two pieces of private information, k_i and c_i, and does not know k_j and c_j.[14] We assume that each firm's priors are that its rival's disclosure costs are independent and normally distributed and that its rival's marginal costs of production are also normally distributed. A key feature of our model is that the covariance between the two firms' marginal costs is non-negative. Thus, each firm can learn about its rival's marginal costs when it learns its own marginal costs. More formally, k_1 and k_2 are independent, normally distributed random variables with zero means and variances $\sigma_{k_1k_1}, \sigma_{k_2k_2}$. Further, $c_i \sim N(\mathrm{E}[c_i], \sigma_{ii})$ for $i = 1,2$ and $\mathrm{Cov}[c_1, c_2] \equiv \sigma_{12} \geq 0$.

A key difference between our model and most others in the literature is that each firm has two sources of private information: its disclosure costs and its production costs.[15] Thus, firm i's disclosure in the first stage is likely to depend on both components of its private information and, therefore, may provide its rival with information about c_i. If so, then the disclosure will affect how the firms compete in the product market.

As is standard, we focus on linear equilibria. In the first stage of the game, each firm knows its own disclosure and production costs. That is, firm i's first period information set is $y_i^1 = (c_i, k_i)$, so we conjecture that each firm uses a strategy that is linear in the elements of its first stage information set. In the second stage, each firm has observed the disclosures made in the first stage and, as a result, firm i's information set is $y_i = (c_i, s_i, s_j)$. While firm i's own disclosure does not provide it with additional information, we include s_i because its rival uses s_i to infer c_i; we do

[12] Quadratic cost functions have become relatively common in the disclosure literature because they support linear equilibria in Cournot-like models. See, for example, Crawford and Sobel (1982), Fischer and Stocken (2001), and Chakraborty and Harbaugh (2007) who employ a quadratic cost function in their models of cheap talk, as well as the models in Stocken and Verrecchia (2004), Fischer and Stocken (2004), and Guttman et al. (2006).

[13] Prior work using a Cournot model with incomplete information about costs include Fried (1984), Shapiro (1986), Gal-Or (1986, 1988), Darrough (1993), and Vives (2002). Also, see the survey by Verrecchia (2001).

[14] Formally, we are assuming that each firm's type is two-dimensional, $t_i = (k_i, c_i) \in \mathcal{R}^2$.

[15] The main exceptions are Fischer and Verrecchia (2000) and studies that build on their model such as Har et al. (2009). In these models, incomplete contracting limits the market's ability to understand how managerial compensation is linked to stock price. Since the manager does know how his/her compensation is linked to stock price and has private information about the firm's future performance, the manager has two sources of private information in Fischer and Verrecchia-type models, although the sources differ from those studied here.

not include k_i because its rival does not observe it and it is independent of all other unknowns. Thus, firm i anticipates that q_j will depend on s_i and therefore conditions its own output choice on s_i. Neither firm's output choice depends on k_i because neither uses its own disclosure to infer its own private information. Consequently, we conjecture that in equilibrium:

$$s_1 = D_0 + D_1 c_1 + D_2 k_1$$

$$s_2 = F_0 + F_1 c_2 + F_2 k_2$$

$$q_1 = N_0 + N_1 c_1 + N_2 s_1 + N_3 s_2$$

$$q_2 = M_0 + M_1 c_2 + M_2 s_1 + M_3 s_2. \tag{C1}$$

To solve for a perfect Bayes equilibrium of the game, we begin by analyzing the second-stage game in which firms compete in the product market given their prior disclosures and their inferences about their rival's private information. Given our normality assumptions and our focus on linear equilibria, we make the additional conjecture that:[16]

$$\mathrm{E}[c_1|y_2] = \alpha_0 + \alpha_1 c_2 + \alpha_2 s_1$$

$$\mathrm{E}[c_2|y_1] = \beta_0 + \beta_1 c_1 + \beta_2 s_2. \tag{C2}$$

Given these conjectures, in the second stage, firm i solves:

$$\max_{q_i} \mathrm{E}[(a - Q)q_i - c_i q_i | y_i] \equiv \max_{q_i} a q_i - q_i^2 - \mathrm{E}[q_j | y_i] q_i - c_i q_i. \tag{1}$$

Note that the first-stage disclosure costs are ignored because, in the second stage, they are a sunk cost and thus do not affect the optimal output decision. The first-order condition for this maximization problem is $0 = a - 2q_i - \mathrm{E}[q_j \mid y_i] - c_i$, which forms the basis for Proposition 1:[17]

Proposition 1: If (C1) and (C2) hold, then there exists a linear equilibrium (q_1^*, q_2^*) to the production game with:

$$q_1^* = N_0 + N_1 c_1 + N_2 s_1 + N_3 s_2$$

$$q_2^* = M_0 + M_1 c_2 + M_2 s_1 + M_3 s_2,$$

where:

$$N_0 = (1/3)\left(a - \frac{(2-\beta_1)\alpha_0}{4-\alpha_1\beta_1} + \frac{2(2-\alpha_1)\beta_0}{4-\alpha_1\beta_1}\right) \qquad M_0 = (1/3)\left(a - \frac{(2-\alpha_1)\beta_0}{4-\alpha_1\beta_1} + \frac{2(2-\beta_1)\alpha_0}{4-\alpha_1\beta_1}\right)$$

$$N_1 = -\frac{2-\beta_1}{4-\alpha_1\beta_1} \qquad M_1 = -\frac{2-\alpha_1}{4-\alpha_1\beta_1}$$

$$N_2 = -(1/3)\alpha_2\left(\frac{2-\beta_1}{4-\alpha_1\beta_1}\right) \qquad M_2 = (2/3)\alpha_2\left(\frac{2-\beta_1}{4-\alpha_1\beta_1}\right)$$

[16] In a linear equilibrium, we will have to compute all of the conjectured coefficients in Equations (C1) and (C2).

[17] The second-order condition is satisfied since the coefficient on q_i is negative.

$$N_3 = (2/3)\beta_2\left(\frac{2-\alpha_1}{4-\alpha_1\beta_1}\right) \qquad M_3 = -(1/3)\beta_2\left(\frac{2-\alpha_1}{4-\alpha_1\beta_1}\right).$$

There are two important features of this linear equilibrium that we should highlight. First, each firm's equilibrium production strategy depends on its rival's disclosure, its actual production costs and its own disclosure. The (potentially) unintuitive feature is that the firm's production strategy depends on its own disclosure despite the fact that the firm knows its own cost of production and uses it rather than its own disclosure to estimate its rival's marginal cost of production. As suggested above, the reason for this dependence is that firm i knows that its rival is using i's disclosure to make inferences about i's production costs (see Equations (C2)) and thus i can infer that its rival's production strategy will depend on s_i. Since firm i's production strategy depends on its inference about firm j's production decision that firm i knows depends on s_i, firm i's production strategy "indirectly" depends on its own disclosure because of the information its rival can extract from that disclosure. Second, the parameters describing the equilibrium production strategies (the M's and N's) depend on the coefficients in the conditional expectations described by (C2). Intuitively, both firms use all of their private information and their public disclosures from the disclosure stage to infer as much as they possibly can about their rival's production costs.

Proposition 1 can also be used to derive each firm's equilibrium expected second stage profits. Substituting q_i^* into Equation (1) shows that, as is usual in Cournot games, firm i's equilibrium expected second stage profits are:

$$\mathrm{E}[\pi_i^2|y_i] = (a - c_i - q_i^* - \mathrm{E}[q_j^*|y_i])q_i^* = (1/2)(q_i^*)^2.$$

It will also be useful to have simplified expressions for $\mathrm{E}[q_j^* \mid y_i]$. This is most readily obtained from Proposition 1 by taking expectations:

$$\mathrm{E}[q_2^*|y_1] = M_0 + M_1\mathrm{E}[c_2|y_1] + M_2 s_1 + M_3 s_2$$

$$\mathrm{E}[q_1^*|y_2] = N_0 + N_1\mathrm{E}[c_1|y_2] + N_2 s_1 + N_3 s_2.$$

Substituting for $\mathrm{E}[c_i \mid y_j]$ from (C2) shows that each equation has the form implied by (C1):

$$\mathrm{E}[q_2^*|y_1] = A_0 + A_1 c_1 + A_2 s_1 + A_3 s_2 \tag{2}$$

$$\mathrm{E}[q_1^*|y_2] = B_0 + B_1 c_2 + B_2 s_1 + B_3 s_2. \tag{3}$$

Folding back to the disclosure stage, each firm chooses a disclosure, s_i, which we interpret as a disclosure about the firm's average (equivalently, in our model, marginal) cost of production. Recall that our interpretation is that the firm's income statement can be used to infer the firm's reported average costs. Given our distributional and information assumptions, in the first stage of the game, firm i chooses s_i to maximize its expected profits in the two-stage game overall:[18]

[18] The maximand also highlights the importance and implications of our assuming quadratic costs of disclosure. In particular, if $\varepsilon_i = 0$ so that disclosure costs are linear, then we will only find corner solutions to the firm's maximization problem. If, on the other hand, $\varepsilon i \neq 0$, then disclosure costs are quadratic, the firm's maximization problem has an interior solution, and we can identify the linear equilibrium we seek. Further, because k_i is the coefficient on the linear term of the disclosure costs, its value only affects the intercept of the firm's reaction function and, thus, the particular value it takes on does not impact our ability to find interior solutions to the firm's maximization problem.

$$\max_{s_i}(1/2)\mathrm{E}[(q_i^*)^2|y_i^1] - k_i(s_i - c_i) - \varepsilon_i(s_i - c_i)^2.$$

Substituting equilibrium outputs from Proposition 1, and maximizing with respect to s_1 (s_2) yields:

$$s_1 = \left[\frac{1}{\varepsilon_1 - A_2^2}\right](-A_2(a - A_0) + (\varepsilon_1 + A_2(1 + A_1))c_1 - k_1 - A_2A_3\mathrm{E}[s_2|c_1,k_1]) \quad (4)$$

$$s_2 = \left[\frac{1}{\varepsilon_2 - B_3^2}\right](-B_3(a - B_0) + (\varepsilon_2 + B_3(1 + B_1))c_2 - k_2 - B_2B_3\mathrm{E}[s_1|c_2,k_2]). \quad (5)$$

Since k_1 and k_2 are not correlated with each other or the firms' marginal costs of production, $\mathrm{E}[s_j \mid k_i, c_i] = \mathrm{E}[s_j \mid c_i]$. Equations (4) and (5) suggest that s_i is linear in c_i, k_i, as long as the conditional expectations $\mathrm{E}[s_j \mid c_i]$ are linear in those variables, too. Proposition 2 provides conditions under which the s_i^*s are, in fact, linear.

Proposition 2: In every linear equilibrium, (q_i^*, q_2^*) are defined as in Proposition 1 and (s_1^*, s_2^*) satisfy:

$$s_1^* = D_0 + D_1c_1 + D_2k_1$$

$$s_2^* = F_0 + F_1c_2 + F_2k_2,$$

where:

$$\mathrm{E}[c_2|c_1] = \mu_0 + \mu_1c_1 \qquad \mathrm{E}[c_1|c_2] = \delta_0 + \delta_1c_2$$

$$D_0 = \frac{1}{\varepsilon_1 - A_2^2}[-A_2(a - A_0 + A_3F_0 + A_3F_1\mu_0)] \qquad F_0 = \frac{1}{\varepsilon_2 - B_3^2}[-B_3(a - B_0 + B_2D_0 - B_2D_1\delta_0)]$$

$$D_1 = \frac{1}{\varepsilon_1 - A_2^2}[A_2(1 + A_1) - A_2A_3F_1\mu_1 + \varepsilon_1] \qquad F_1 = \frac{1}{\varepsilon_2 - B_3^2}[B_3(1 + B_1) - B_2B_3D_1\delta_1 + \varepsilon_2]$$

$$D_2 = -\left(\frac{1}{\varepsilon_1 - A_2^2}\right) \qquad F_2 = -\left(\frac{1}{\varepsilon_2 - B_3^2}\right).$$

We also observe that in any linear equilibrium as described in Proposition 2:

$$\mathrm{E}[c_2|c_1,s_2] = \mathrm{E}[c_2] + \left(\frac{\sigma_{12}\sigma_{k_2k_2}F_2^2}{\sigma_{11}(\sigma_{22}F_1^2 + \sigma_{k_2k_2}F_2^2) - \sigma_{12}^2F_1^2}\right)(c_1 - \mathrm{E}[c_1])$$
$$+ \left(\frac{\sigma_{11}\sigma_{22}F_1 - \sigma_{12}^2F_1}{\sigma_{11}(\sigma_{22}F_1^2 + \sigma_{k_2k_2}F_2^2) - \sigma_{12}^2F_1^2}\right)(s_2 - F_0 - F_1\mathrm{E}[c_2])$$

$$\mathrm{E}[c_1|c_2,s_1] = \mathrm{E}[c_1] + \left(\frac{\sigma_{12}\sigma_{k_1k_1}D_2^2}{(\sigma_{11}D_1^2 + \sigma_{k_1k_1}D_2^2)\sigma_{22} - \sigma_{12}^2D_1^2}\right)(c_2 - \mathrm{E}[c_2])$$
$$+ \left(\frac{\sigma_{11}\sigma_{22}D_1 - \sigma_{12}^2D_1}{(\sigma_{11}D_1^2 + \sigma_{k_1k_1}D_2^2)\sigma_{22} - \sigma_{12}^2D_1^2}\right)(s_1 - D_0 - D_1\mathrm{E}[c_1]),$$

and a linear equilibrium requires that these coefficients match the α's and β's in Equations (C2).

Unfortunately, there is no general result on the existence of a perfect Bayes equilibrium in this type of linear-normal model.[19] Further, finding a linear equilibrium in closed form for our model is very difficult because it requires solving a system of 24 nonlinear equations for the unknown coefficients. Both numerical solutions and closed-form solutions for restricted versions that eliminate some of the key forces in the model are analyzed below. However, before examining the properties of the solutions, we explain the importance of our assumption that there are two sources of private information.

The easiest way to see the importance of assuming that there are two sources of private information rather than just one is to consider the effect of assuming that both firms know the realizations of k_1 and k_2 (equivalently that $\sigma_{k_1k_1} = \sigma_{k_2,k_2} = 0$). If so, then each firm has one source of private information: their own costs of production. As a result, any disclosure strategy that is monotone in the firm's private information can be inverted, which allows all to infer the firm's private information exactly and we are able to prove the following result:

Proposition 3: If k_1 and k_2 are common knowledge, then there are no monotone equilibria in pure strategies.

The key implication of assuming that k_1 and k_2 are common knowledge is that neither firm can employ a monotone disclosure strategy without revealing its actual production costs. That is, if the conjectured disclosures are monotone in the firm's private information, even if biased, then they allow each firm to infer the other's costs and thus be fully informed when they compete in the product market. As a result, if a firm deviates from the conjectured disclosure strategy and biases its reported costs downward, then the rival would infer that the deviating firm's costs are lower and accommodate the firm by producing less. This allows the deviating firm to produce more and shift product market profits from its rival to itself.

In the following section, we examine restricted versions of our model that allow for closed-form solutions at the cost of eliminating key components of the information extraction problem and how rivalry impacts the firms' decisions to bias their reported earnings. In Section IV, we examine numerical solutions of the full model and highlight both of these features.

III. RESTRICTED VERSIONS OF THE MODEL

In this section, we examine two special cases of the model described above that simplify the computation of a linear equilibrium at the cost of eliminating important forces in the model.

We begin by restricting the analysis to the case when the firms' private information is uncorrelated. As we noted in the previous section, this restriction eliminates a key force in our model—the impact of learning about what the rival can learn from the firm's own financial reports. This assumption, that $\sigma_{12} = 0$, implies that, in Proposition 2, α_1 and β_1 are both zero. While it is true that eliminating this dimension of each firm learning about the information its rival can extract from its financial reports does simplify the system of equations mentioned above, it does not simplify them enough to offer useful closed-form solutions.[20]

To obtain a manageable linear equilibrium, we introduce a second restriction: only one firm has private information. The cost of this restriction is that it eliminates the impact of rivalry on the

[19] With minor modifications to our distributional assumptions, however, the existence of a trembling hand perfect Nash equilibrium ensures that there is a perfect Bayes equilibrium (see Fudenberg and Tirole 1993). In particular, if each firm's private information is drawn from a finite set (so that the set of types in our game is finite), standard existence theorems ensure that there is a trembling hand perfect Nash equilibrium and therefore a perfect Bayes equilibrium in the modified game.

[20] Algebraic manipulation, available from the authors, produces a quadratic equation with one real root. Unfortunately, the root is a very complex function of the exogenous variables and requires that we again revert to numerical solutions.

earnings management decision as well as a key component of each firm's inference problem. Without loss of generality, assume that c_2 is common knowledge and since firm 2's costs are known, set $k_2 \equiv 0$.

Corollary: If c_2 is common knowledge, $k_2 \equiv 0$ and $\sigma_{12} = 0$, then there is a linear equilibrium $\{(s_1^*, q_1^*(s_1, s_2)), (s_2^*, q_2^*(s_1, s_2))\}$ defined by:

$$q_1^* = (1/6)(2(a + \beta_0) - 3c_1 - s_1 + 2\beta_2 s_2)$$

$$q_2^* = (1/6)(2(a - (1/2)\beta_0) - 3c_2 + 2s_1 - \beta_2 s_2)$$

$$s_1^* = \left(\frac{1}{9\varepsilon_1 - 1}\right)[2a - (1/2)\beta_0 + (3 + \varepsilon_1)c_1 + (1/2)(3 + \beta_2)c_2 - 9k_1]$$

$$s_2^* = c_2,$$

with:

$$\mathrm{E}[c_1|c_2, s_1] = \beta_0 + \beta_2 s_1.$$

$$\mathrm{E}[c_2|c_1, s_2] = c_2$$

$$\beta_0 = \left(1 - \frac{(3 + \varepsilon_1)\beta_2}{9\varepsilon_1 - 1}\right)\mathrm{E}[c_1]$$

$$\beta_2 = \left(\frac{1}{9\varepsilon_1 - 1}\right)^2\left[\frac{(9\varepsilon_1 - 1)(3 + \varepsilon_1)\sigma_{11}}{(3 + \varepsilon_1)^2\sigma_{11} + 81\sigma_{kk}}\right].$$

Despite the significant restrictions that eliminate many key forces, the equilibrium described in the Corollary does have a couple of interesting features. First, because firm 2's costs are common knowledge, only firm 1 has the opportunity to impact product market competition by its disclosure policy—and it chooses to do so in equilibrium. The reason it chooses to do so is that firm 2 cannot "look through" firm 1's disclosure and correctly infer firm 1's costs of production. The two sources of private information mean that firm 2 optimally attributes part of the difference between firm 1's reported cost (s_1^*) and its expected cost ($\mathrm{E}[c_1]$) to the realized value of k_1 and the rest to firm 1's cost of production (c_1). Thus, firm 1 has an incentive to underreport its costs (when $k_1 = 0$) and, despite firm 2 optimally updating its expectation of firm 1's costs, firm 2 is left with the belief that firm 1's costs are lower than they actually are. Firm 2, therefore, optimally responds to its belief about firm 1's costs by producing fewer units than it would if it knew firm 1's costs. Firm 1 anticipates this reaction to its disclosure and produces more than it would had firm 2 known firm 1's costs. The end result is that profits are shifted from firm 2 to firm 1 providing the motivation for firm 1 to bias its disclosure. Put differently, competition in the product market induces firm 1 to bias its report and the biased report alters the competitive outcome in the product market.

A disadvantage of this restricted model is that there are few interesting comparative statics. As one would expect, as firm 1's cost of providing a misleading disclosure increases, it offers a less biased report and earns smaller profits. For essentially the same reasons, if the size of the product market is smaller (α declines), then firm 1 offers a less biased report and earns smaller profits. In the first case, the cost of biasing its disclosure increases and so it does less. In the latter case, the benefits of biasing its disclosure decrease and so it does less.

Interestingly, despite the limitations of this version of the model, it still contributes to the existing literature on biased reporting. For example, Fischer and Verrecchia (2004) study transparent disclosure bias in a Cournot oligopoly model with firms that employ heuristics to interpret new information. In their model, some firms over-react to information and, in equilibrium, the informed firm biases its disclosure. In essence, firms that employ heuristics "believe" their disclosure/interpretation and produce more output than they would if they were Bayesian. Bayesian firms realize this and accommodate the added output, thereby producing a benefit to the heuristic firms. Our restricted model complements Fischer and Verrecchia's (2004) analysis by showing that biased reporting arises even when every firm optimally employs all available information (every firm is Bayesian) but still produces a similar shifting of product market profits. In contrast to Fischer and Verrecchia (2004), the bias is not fully transparent and so our analysis suggests that fully transparent biased reporting may require the non-optimal use of available information (heuristics).

Our model also complements Fischer and Verrecchia's (2000) analysis of reporting bias. In their model, a manager is permitted to offer biased earnings reports. They show that, when investors are uncertain about how the manager's compensation depends on stock price, in equilibrium, the manager offers a biased voluntary disclosure that is not transparent. In their model, the market's uncertainty about the manager's payoff is required for equilibrium to exhibit biased reports. Our model complements their analysis by showing that competition in the product market can substitute for agency problems combined with incomplete contracting as an explanation for biased reporting.

Our results for the restricted model differ markedly from those reported in the next section when both key features of our model—that both firms face an information-extraction problem and that there is an important component of rivalry in determining how each firm manages its earnings (in addition to the rivalry associated with the firms' competition in the product market)—are present. In the full model, both firms bias their reports and this allows both to earn greater profits than they would in the full-information environment. Further, the restrictions do not allow us to explore how earnings management is affected by how much information a firm can infer about its rival or how asymmetries in the costs of misleading disclosures impacts equilibrium outcomes.

IV. SYMMETRIC EQUILIBRIA

In this section, we focus on symmetric equilibria, and in the following section, we explore the nature of asymmetric equilibria. In a symmetric equilibrium, both firms employ the same strategies (the same map from their private information to their actions). Thus, equilibrium disclosures are determined by the same function of the firm's private information, and equilibrium production decisions are determined by the same function of each firm's information set in the second stage. That is, symmetry requires that $D_i = F_i$ and $\alpha_i = \beta_i$ for $i = 1,2,3$; $N_i = M_i$ for $i = 1,2,3,4$ and $\mu_i = \delta_i$ for $i = 1,2$. Further, we must assume that both firms are in the same "competitive position:" k_1 and k_2 must be drawn from the same distribution, c_1 and c_2 must be drawn from the same distribution and the firms must incur the same disclosure costs, $\varepsilon_1 = \varepsilon_2 \equiv \varepsilon$. Finally, without loss of generality, we assume that all of the normally distributed random variables are standard normal with zero means and variances equal to 1. Given this structure, we have:

Proposition 4: If ε is not too small, then there is a symmetric linear perfect Bayes equilibrium in pure strategies.

Intuitively, if the cost of misreporting is sufficiently small, there are no linear equilibria because the incentive to misreport is too great and cannot depend linearly on the firm's private information. However, if the cost of misreporting is larger, then there is a linear perfect Bayes equilibrium in which each firm provides a disclosure. The disclosure allows the rival to update its

beliefs about the disclosing firm's production costs but does *not* allow it to infer exactly what those costs are (i.e., the biased reports are not transparent). The update does, however, affect how the rival competes with the disclosing firm in the product market.

To better understand the properties of the disclosure and its effects on product market competition, we will generally do numerical comparative static analysis because of the complexity of the equations defining the symmetric equilibrium. However, comparative static results on the cost of misreporting can be done analytically. In particular, as the cost of misreporting increases, the amount of misreporting declines, each firm's estimate of its rival's cost of production becomes more accurate, and each firm's output approaches the output it would make in a full-information environment.[21]

Intuitively, this comparative static tells us that, as the cost of misreporting gets large, equilibrium in the product market converges to the complete information solution. The reason is that, as the cost of misreporting gets large, the firms do less which makes their disclosures more informative. Thus, the uncertainty about the rival's cost dissipates and competition in the product market approaches the full-information outcome. Interestingly, this result does not require that firms precommit to a disclosure or be required to disclose truthfully. Instead, as the cost of misreporting rises, firm behavior converges to the full-information solution where each truthfully discloses its private information about its production costs.

To clarify how equilibrium outcomes vary with the cost of misreporting, the correlation between the firm's costs of production and the size of the product market, we turn to numerical comparative static analyses. Since $h_i(s_i, c_i) = k_i(s_i - c_i) + (1/2)\varepsilon_i(s_i - c_i)^2$, a positive (negative) k_i produces marginal incentives for the firm to report $s_i < c_i$ ($s_i > c_i$). Thus, we do all numerical analyses assuming that the realized values of k_1 and k_2 are zero. We begin by analyzing the equilibrium described above in more detail and the effects of an increase in the cost of misreporting.

Result 1: In equilibrium, each firm's reported cost of production is smaller than its actual cost of production, each firm's output is smaller than the full-information quantity and each firm's estimate of its rival's costs is smaller than its rival's actual cost of production.[22]

This result tells us that firms bias their reported costs downward in an attempt to create the impression that their costs are lower than they actually are.[23] If successful, their rival will optimally reduce the amount it sells in the product market, causing the price to be higher and resulting in greater profits for the misreporting firm. Each firm, in equilibrium, is successful because neither can perfectly extract its rival's production costs from reported costs. Each understands the equilibrium and adjusts its expectation of the rival's cost taking into account the rival's incentive to report lower costs but, despite this, the adjustment is only partial. We emphasize that neither firm is "fooled." Each uses all available information (its private information about its own costs and the

---

[21] See the Appendix for a proof of these assertions.

[22] Further, combining Result 1 with our comparative static results on changes in ε, when the cost of misreporting increases, all differences shrink. Reported costs converge to actual costs, output choices converge to the full-information quantities, and each firm's estimate of its rival's costs converge to its rival's actual costs.

[23] We should clarify that we are not using biased in the classical statistical sense but rather, to use Fischer and Verrecchia's (2000) term, in the "realized" sense. This distinction is standard in the biased disclosure literature and is important to understanding our results because both firms are Bayesian, which implies that there is no bias in the classical statistical sense. There is, however, bias in the "realized" sense that can be seen as follows. Since s_i is linear in c_i and k_i, it is normally distributed and so, after suppressing conditioning on c_j, we can write $E[c_i \mid s_i] = E[c_i] + \kappa(s_i - E[s_i])$. Recalling that $k = 0$, substituting for s_i and $E[s_i]$ and rearranging, we obtain $\lambda c_i - E[c_i \mid s_i] = -(1-\lambda)E[c_i]$. Evaluating at $E[c_i] = 0$ and noting that $0 < \lambda < 1$, we have that $c_i > E[c_i \mid s_i]$. That is, each firm's estimate of its rival's costs is smaller than its rival's actual cost of production.

disclosures) to make the best possible estimate of its rival's production costs. Each underestimates its rival's costs because there are two reasons why they can be small—either the costs of production actually are small; or the costs of misreporting are small, making it less expensive to misreport—and the firm cannot completely sort out the two effects.[24]

These results differ markedly from the results for Cournot models in the prior literature (Darrough 1993; Raith 1996). Darrough (1993) proves that, in a Cournot model, if the firm's disclosure is required to be a noisy but unbiased reflection of the firm's private information about its cost of production, then in equilibrium, both firms would precommit to minimize the noise in the disclosure.[25] That is, they report their private information without bias. Other models permit firms to choose whether to disclose after learning their production costs but require truthful disclosures. This literature, summarized in Christensen and Feltham (2002), shows that the standard unraveling result applies—in equilibrium, every firm-type discloses because if more than one type is pooled together, then the type with the lowest production costs can increase profits by disclosing its costs and inducing its rivals to reduce the quantities they offer for sale. In our model, firms are not required to disclose truthfully, but if they report with a bias, they incur a cost associated with misreporting. Our analysis shows that when firms have the option to bias their disclosure (even at a cost), they do so and that the effects are only partially accounted for by the rival.

Result 1 also highlights the importance of rivalry in the earnings management decision.[26] In particular, even though both firms are fully rational and optimally use all available information, biased reporting occurs in equilibrium and leads to greater industry profits. This differs from both the results in Fischer and Verrecchia (2004) and our restricted model. In particular, in Fischer and Verrecchia (2004), profits are shifted from firms that use Bayes' Rule to those using a particular heuristic and leads to lower overall industry profits. In our restricted model, profits are shifted from the firm without private information to the firm with private information (via the latter's biased disclosure) and again leads to lower industry profits. In contrast to the aforementioned models without rivalry in the disclosure stage, when there is rivalry, both firms earn greater profits than the full-information outcome because both reduce output after optimally using the rival's disclosure to estimate their rival's cost of production. The impact on firm profits is described in the following result:

> **Result 2:** In equilibrium, misreporting allows both firms to earn greater profits in the product market than they would if each had complete information about its rival's cost of production. Further, the profits the firms earn in the product market decline as the cost of misreporting increases.

Intuitively, the type of earnings management we examine allows each firm to bias its disclosure in order to affect its rival's inference about its production costs. Result 1 indicates that the

---

[24] The direction of the preferred bias (to report production costs as being smaller than they actually are) is the result of assuming that the firms are Cournot competitors. Had we assumed that the firms sold heterogeneous products and were Bertrand competitors, the direction of bias would be reversed. Formally, the reason for this difference is that quantities are strategic substitutes whereas prices are strategic complements. Intuitively, the reason is that under Bertrand competition, if firm i reports higher costs of production, it anticipates that its rival will increase its price, allowing i to increase its price thereby increasing i's profits. The higher price increases firm profits because competition has driven the equilibrium prices below their monopoly levels.

[25] Formally, if the firm's private information is ψ_i, then it is required to disclose $\hat{\psi}_i = \psi_i + \nu_i$ where ν_i is a normally distributed random variable with mean 0 and variance chosen by the disclosing firm. If the firm wishes to disclose its private information, then it selects a variance of zero and if it wishes not to disclose, it selects a variance of infinity. In neither case can the disclosing firm intentionally bias its disclosure.

[26] Recall that rivalry requires two assumptions: (1) that at least two firms have private information, and (2) that their information is correlated. Thus, the impact of rivalry in our model is seen by contrasting equilibrium outcomes in the restricted model (Corollary) and in our full model.

firms choose to misreport costs to create the impression that they are lower than they actually are. As a result, the rival infers that the firm will produce more than it actually does and responds by reducing the quantity it produces. Consequently, earnings management results in both firms selling less than they would, had they known their rival's costs, but the associated price increase (from selling less) results in each firm earning greater profits than it would with complete information about its rival's costs. Further, as the cost of misreporting rises, the magnitude of the bias in the firms' reports declines. Since both firms' profits in the product market are increasing in the amount of earnings management, increases in the cost of misreporting lower the equilibrium amount of earnings management and therefore lower the profits each firm earns in the product market. Further, if the costs of misreporting are nonpecuniary or are incurred by management rather than the firm, Result 2 implies that earnings management actually increases the liquidation value of the firm.

The clearest empirical implications follow from associating changes in the cost of misreporting with changes in reporting regulations. For example, Section 302 of the Sarbanes-Oxley Act of 2002 (SOX) requires, among other things, that CEOs and CFOs (or persons performing equivalent functions) personally certify in each quarterly and annual report, including transition reports, that:

> he or she has reviewed the report; based on his or her knowledge, the report does not contain any untrue statement of a material fact or omit to state a material fact necessary in order to make the statements made, in light of the circumstances under which such statements were made, not misleading with respect to the period covered by the report; based on his or her knowledge, the financial statements, and other financial information included in the report, fairly present in all material respects the financial condition, results of operations and cash flows of the issuer as of, and for, the periods presented in the report.[27]

Such certification increases the costs of misreporting and would be represented by an increase in ε in our model. Thus, the results on increases in ε and Result 1 combine to suggest the intuitive result that the impact of Section 302 is to reduce the amount of misreporting and the surprising results that: (1) each firm's estimate of its rival's costs increases and becomes more accurate, and (2) each firm produces additional output. However, to the extent that misreporting of the type we consider is important to the economy, our analysis suggests that Section 302 of SOX likely expanded gross domestic product but reduced taxable profits and tax revenues.

More importantly, our analysis suggests that only certain types of misreporting are effective means of creating a competitive advantage in the product market. In particular, aggressive cost capitalization, including operating costs in restructuring costs, or selling previously written off inventory all produce financial statements that lead the firm's rival to infer that the firm's costs are lower than they actually are. Changes in certain estimates can also produce the same result. For example, a firm that reduces the allowance for doubtful accounts (to increase earnings rather than because of changes in its customers' credit-worthiness) increases reported revenue relative to operating expenses, and thus gives its rival the impression that the reporting firm's costs are lower than they actually are. Another example is reducing warranty expense estimates (without an associated change in the reliability of the product) that directly lowers reported operating expenses and misleads the firm's rival in a similar manner. Finally, certain methods of "real" earnings management such as delaying expenditures have the same effect.

There are, however, a variety of "standard" earnings management techniques that do not create a competitive advantage of the type we examine. For example, aggressive revenue recog-

[27] See "Certification of Disclosure in Companies' Quarterly and Annual Reports," Securities and Exchange Commission 17 CFR PARTS 228, 229, 232, 240, 249, 270 and 274 [RELEASE NOS. 33-8124, 34-46427, IC-25722; File No. S7-21-02] RIN 3235-AI54, http://www.sec.gov/rules/final/33-8124.htm.

nition or granting lenient credit terms (or, more aggressively, channel stuffing) all produce greater earnings (assuming gross margins are positive), but do not necessarily affect a rival's ability to use the firm's financial statements to infer its costs. Similarly, delays in writing down assets, over-reserving for contingencies or "timely" selling of assets are all means of managing earnings that do not alter the rival's ability to infer costs. Thus, assuming that firms generally face similar incentives to engage in this second class of earnings management techniques, our analysis suggests that SOX will reduce the use of the first class of earnings management techniques relative to the second. As a result, our model predicts that the observed proportion of the first class of earnings management techniques among SEC enforcement actions will be smaller post-SOX.

Result 2 also offers an alternative explanation for the "contagion" effect associated with restatements studied recently by Gleason et al. (2008), Xu et al. (2006), Da Dalt and Margetis (2004), Margetis (2004), and Gonen (2003). These studies all examine how a restatement by one firm affects the stock price of its peers. They generally find that the negative effect on the restating firm's stock price is contagious in the sense that the restatement is associated with negative effects on the stock price of its peers. These studies generally focus on the likely transmission mechanism being the market's concern that the restatement by one firm increases the likelihood of a restatement by its peers (hence the name "contagion effect"). Our analysis offers a complementary transmission mechanism for a subset of the financial restatements examined: Result 2 tells us that if the restatement by one firm increases the cost of misreporting, then each firm will respond by reducing the amount of misreporting, thereby lowering the profits each earns in the product market. This complementary transmission mechanism can be examined empirically by focusing on the future operating performance of the restating firm's peers. A deterioration in this performance would be consistent with a loss of profitability in the firms' product market and would suggest that the observed contagion effect may partly reflect the market's anticipation of this future deterioration in performance.[28]

Having analyzed the impact of earnings management on competition in the product market, we can expand our understanding of the incentives to manage earnings by examining how changes in the product market affect each firm's willingness to misreport costs.

Result 3: When firms compete in more profitable product markets, the magnitude of misreporting is greater but the bias in the rival's estimate of a firm's cost of production is smaller, which, in turn, results in a smaller gain in equilibrium product market profits relative to the full-information level of profit.

Intuitively, Result 3 shows how the competitive environment in the product market affects incentives to engage in the type of earnings management we analyze. In particular, firms that compete in more profitable product markets (markets with larger demand intercepts) bias their reported earnings more because the benefits from inducing a rival to reduce output are greater. What may be less obvious is that the induced error in the rival's expectation of the firm's cost is smaller. The reason is that the per-unit cost of misreporting is constant so that as the profitability of the product market increases, inducing firms to report smaller costs, the total cost of misreport-

[28] Recently, Durnev and Mangen (2009) and Kedia and Philippon (2009) have offered alternative but similar explanations for the contagion effect. Durnev and Mangen (2009) suggest a transmission mechanism that relies on prior misstatements causing inefficient investments by a firm's peers. Based on this idea, they offer empirical evidence suggesting that a restatement by one firm in the industry is associated with more inefficient investment decisions by its rivals as measured by Tobin's marginal *q*. Kedia and Philippon (2009) provide a signaling explanation for why fraudulent accounting can affect not only the firm's peers but also have economy-wide effects when the equilibrium is characterized by a pooling equilibrium. Intuitively, in their model, mimicking behavior by low-productivity firms requires that they overinvest and overhire and both effects are reduced following the market's recognition that there has been significant accounting fraud by some firms in the economy.

ing rises. Thus, the increase in the amount of misreporting is tempered by the increasing total costs of misreporting, which, in net, allows its rival to better estimate the firm's true cost of production. Since the estimate is better, the impact of misreporting on the rival's output choice declines and the equilibrium outcome converges to the full-information solution.

An immediate implication of Result 3 is that reported (actual) earnings are positively (negatively) correlated with the magnitude of misreporting. As a result, our analysis predicts that most measures of profitability (e.g., Return on Equity, Return on Invested Capital, Return on Sales) calculated using the firm's unadjusted (adjusted) financial statements should be positively (negatively) correlated with the use of the types of earnings management techniques we analyze. In many cases, this result may be difficult to test because adjusted financial statements are not available. However, this difficulty might be avoided for the subset of firms that voluntarily or involuntarily provide a restatement of past financial information. For these firms, unadjusted and adjusted measures of profitability are available as are the reasons for the restatement. Together, this information would allow for an empirical test of our result. Alternatively, Result 3 implies that firms in highly profitable markets should be those firms that are more often engaged in (or engage in more) misreporting than firms in less profitable markets. As a result, our analysis suggests that firms in more profitable markets are more likely to be represented in the set of firms identified by SEC enforcement actions as having engaged in the types of earnings management that result in biasing expectations about the reporting firm's costs of production.

Result 4: The more information a firm can extract about its rival's cost from knowing its own cost, the smaller is the magnitude of the misreporting, the smaller is the induced bias in the rival's estimate of the firm's cost, and the smaller is the deviation from the full-information level of output.

The amount of information a firm can extract about its rival's costs from knowing its own costs is "measured" by the covariance between the costs, σ_{12}. In particular, the greater the covariance, the greater is the reduction in the firm's remaining uncertainty about its rival's cost after observing its own costs. The reduction in this uncertainty reduces the effectiveness of misreporting by offering a biased report, and the result is that the product market equilibrium is closer to the full-information outcome than it would be, had the firms' costs been less highly correlated.

Result 4 suggests that firms with similar technologies, such as those controlled by physical or chemical processes, should be expected to be less likely to engage in misreporting, and if they do so, to do so in smaller amounts than firms whose technologies are more likely to be very different (e.g., service industries). We also expect that firms in mature industries such as the auto or steel industries are less likely to engage in (or to do relatively less) misreporting because they are likely to have more information about their rival's production costs. Similarly, firms with large portfolios of products are more likely to engage in misreporting because their rivals are likely to be less able to infer the firm's costs of producing each product. This result could be empirically tested by relating misreporting to the number of business segments in the firm.

In summary, Results 1, 2, 3, and 4 show the full effect of firms exercising the option to bias their reported costs. Results 1 and 2 show how a firm's decision to bias its disclosure impacts competition in its product market, and Results 3 and 4 show the feedback effect—how the competitive environment in the product market affects incentives to manage earnings.

V. ASYMMETRIC EQUILIBRIA

The results in the previous section provide new insights into the relation between product market competition and misreporting designed to bias a rival's inferences about the reporting firm's cost of production. The main feature that cannot be explored when the focus is on symmetric equilibria is the impact of differential costs of misreporting.

Previously, the cost of misreporting was treated as either non-random or as the expected cost of misreporting.[29] The advantage of interpreting it as an expected cost is that it is natural to assume that firms face different expected costs of misreporting. Not only might firms differ in the likelihood that they are examined for misreporting, but it is also possible that firms are run by managers who differ in their concern about the impact of being detected misreporting. This situation appears to apply in the AT&T-WorldCom example cited in the introduction.

To explore the effect of different costs of misreporting, we examine asymmetric equilibria. We again apply numerical techniques because of the difficulties solving the large system of nonlinear equations that describe the equilibrium. The key difference in this section is that we relax the assumption that $\varepsilon_1 = \varepsilon_2$ by assuming, without loss of generality, that $\varepsilon_2 > \varepsilon_1$. As in the prior section, all numerical analyses assume that the realized values of k_1 and k_2 are zero.

Result 5: When ε_1 and ε_2 are sufficiently different, as ε_2 rises relative to ε_1:

(i) the amount of misreporting by the low-cost misreporter increases while the amount of misreporting by the high-cost misreporter declines;
(ii) the bias in the estimate of the low-cost misreporter is greater but both converge to zero as ε_2 rises;
(iii) the low-cost misreporter produces more output than the full-information output, while the high-cost misreporter produces less; and
(iv) the low-cost misreporter earns greater product market profits than the high-cost misreporter does and the difference increases as ε_2 increases.

Result 5 describes how differences in the costs of misreporting affect competition. In particular, the greater the difference in the costs of misreporting, the larger (smaller) is the amount done by the low-cost (high-cost) misreporter. Thus, holding the lower cost of misreporting constant, Result 5 suggests that increases in the high-cost firm's costs motivate its rival to engage in more misreporting. In fact, Result 5 says even more: If the low cost rises slower than the high cost, then the direct effect of the cost increase on the low-cost misreporter (to reduce the amount of misreporting) is overwhelmed by the indirect, competitive effect of the increase in its rival's cost of misreporting. Further, the lower cost of misreporting offers the firm a competitive advantage in the product market, which it exploits by selling more output than its rival and earning greater profits.

Result 5 also suggests some empirically testable implications. First, if firms "forced" to replace upper management in the wake of a misreporting scandal incur higher costs of misreporting, then rivals are likely to engage in more misreporting and increase production, thereby shifting profits away from the firm that changed upper level management. Second, it suggests a potential risk from using deferred prosecution agreements in misreporting cases.[30] Such agreements increase (potentially dramatically) the cost of misreporting, and our Result 5 suggests that such agreements will produce unwanted competitive responses from the firm's rivals. In particular, the agreements will motivate the firm's rivals to increase their level of misreporting.

VI. CONCLUSION

We examine how earnings management affects product market competition and how rivalry impacts both production decisions and disclosure (earnings management) decisions. To do so, we

[29] If we assume that the cost of misreporting is uncorrelated with the firm's private information, then because the firm's payoffs are linear in that cost, we could take expectations with respect to it and simply replace the random variable with its expectation in the expressions derived above.

[30] Examples of the use of such agreements regarding misreporting of the type we examine include the Bristol-Myers Squibb case (http://www.usdoj.gov/usao/nj/press/files/pdffiles/deferredpros.pdf); the PNC Financial case (http://www.usdoj.gov/opa/pr/2003/June/03_crm_329.htm); the Computer Associates case (http://www.sec.gov/litigation/litreleases/lr18891.htm); and the AOL case (http://www.usdoj.gov/opa/pr/2004/December/04_crm_790.htm).

consider an incomplete information Cournot duopoly model in which firms know their own production costs but not their rival's. In our model, fully rational firms provide a disclosure (e.g., an income statement) that its rival can use to update its beliefs about the disclosing firm's production costs prior to competing in the product market. Our model differs from those in the prior literature in that we allow firms to provide biased reports, but if they do so, they incur a cost of misreporting. We show that biased reporting is impossible in monotone equilibria when these costs are common knowledge. However, when costs are private information, we find a linear equilibrium in which both firms bias their reports even though both are fully rational and optimally use all information when estimating their rival's costs of production. Thus, our model complements and extends Fischer and Verrecchia (2004) by showing that biased reporting does not require firms to employ heuristics rather than optimally update their beliefs. Further, our model complements and extends Fischer and Verrecchia (2000) by showing that agency problems and incomplete contracting are not required for equilibrium to exhibit biased disclosures. Finally, we extend the literature on disclosure in oligopoly by allowing biased disclosure and showing that, when it is available, firms choose to bias their disclosures.[31]

Intuitively, in our model, firms have an incentive to bias their reported costs downward and, even though all firms use all available information efficiently, fully understand each other's incentives, and adjust their beliefs about their rival's costs upward, they still underestimate those costs. As a result, each cuts production relative to the full-information level and each earns greater product market profits. Interestingly, these effects are smaller in more profitable product markets even though the magnitude of misreporting increases. They are also smaller when firms use more similar production technologies. Finally, when firms have sufficiently different (expected) costs of misreporting, the bias in the low-cost misreporter's disclosure is greater than its rival's and it produces more output than the full-information quantity.

By focusing on competitive effects and fully rational firms, our analysis suggests some empirical implications. To see why, note that aggressive cost capitalization, fraudulent revenue recognition, inappropriate estimates of the allowance for doubtful accounts, or warranty expense are all examples of (one class of) earnings management techniques that lead to under-estimates of the firm's production costs, whereas other standard earnings management techniques (the second class) such as channel stuffing, delayed write-downs of assets, or "timely" sales of assets do not. Since incentives to use the second class of earnings management techniques are likely to be independent of the incentives to employ the first class, our model predicts that regulatory increases in the cost of misreporting (e.g., Section 302 of the Sarbanes-Oxley Act of 2002) will result in a change in the distribution of observed earnings management techniques (as, for example, among SEC enforcement actions). In particular, we expect a reduction in the use of the first class of earnings management techniques relative to the second. Second, our analysis suggests that standard measures of profitability will be positively correlated with the use of the first class of earnings management techniques. Third, firms with more similar technologies (e.g., those for whom production is governed by physical or chemical processes or those used in mature industries) will be less likely to employ this type of earnings management than firms with less similar technologies (e.g., service industries or firms that produce a large variety of different products). In addition, our analysis also suggests that the more frequent use of deferred prosecution agreements may lead to unexpected consequences—providing the firm's competitors with incentives to misreport so as to bias estimates of their production costs downward.

Our analysis also appears to support the claims made by C. Michael Armstrong, former CEO of AT&T, that AT&T's *perceived* strategic failures, its inability to compete with WorldCom, and

[31] See Darrough (1993) or Raith (1996) among others or the excellent survey by Verrecchia (2001).

the decision to break up the company were the result of accounting fraud at WorldCom. He suggests that WorldCom's "revenues were false, margins were false, their costs were false" (McConnell et al. 2002, 1) and that this resulted in layoffs, cost cutting, and, finally, the decision to break up AT&T in order to service its debt. Former Sprint CEO William Estry suggests that his company also struggled with its inability to match WorldCom's performance and noted that: "It never dawned on us the base of their pricing was fraud" (Searcy 2005, C1). Our analyses, particularly when one firm's costs of misreporting are greater than its rival's, are consistent with Armstrong's assertions about the impact of WorldCom's accounting fraud on its competitors.

Finally, we should comment on the impact of our assuming that firms are Cournot competitors. It is well known (e.g., Darrough 1993; Raith 1996) that results may be sensitive to the form of competition assumed because quantity choices are strategic substitutes but price choices are strategic complements. We believe that the only significant impact of assuming Bertrand (price-choosing) rather than Cournot (quantity-choosing) competition in our setting is on the direction of the equilibrium bias in reports. Because prices are strategic complements under Bertrand competition, if the firms choose prices and sell substitute products, then we expect that each would have an incentive to bias reported costs of production up rather than down. Other than the direction of the bias, we believe that the remaining results would be qualitatively similar to those obtained under Cournot competition.

APPENDIX

Proposition 1

If (C1) holds, then there exists a linear equilibrium (q_1^*, q_2^*) to the production game with:

$$q_1^* = N_0 + N_1 c_1 + N_2 s_1 + N_3 s_2$$

$$q_2^* = M_0 + M_1 c_2 + M_2 s_1 + M_3 s_2,$$

where:

$$N_0 = (1/3)\left(a - \frac{(2-\beta_1)\alpha_0}{4-\alpha_1\beta_1} + \frac{2(2-\alpha_1)\beta_0}{4-\alpha_1\beta_1}\right) \qquad M_0 = (1/3)\left(a - \frac{(2-\alpha_1)\beta_0}{4-\alpha_1\beta_1} + \frac{2(2-\beta_1)\alpha_0}{4-\alpha_1\beta_1}\right)$$

$$N_1 = -\frac{2-\beta_1}{4-\alpha_1\beta_1} \qquad M_1 = -\frac{2-\alpha_1}{4-\alpha_1\beta_1}$$

$$N_2 = -(1/3)\alpha_2\left(\frac{2-\beta_1}{4-\alpha_1\beta_1}\right) \qquad M_2 = (2/3)\alpha_2\left(\frac{2-\beta_1}{4-\alpha_1\beta_1}\right)$$

$$N_3 = (2/3)\beta_2\left(\frac{2-\alpha_1}{4-\alpha_1\beta_1}\right) \qquad M_3 = -(1/3)\beta_2\left(\frac{2-\alpha_1}{4-\alpha_1\beta_1}\right).$$

Proof

To find a linear equilibrium, compute using the conjectured equilibrium strategies:

$$\mathrm{E}[q_2|y_1] = M_0 + M_1\mathrm{E}[c_2|y_1] + M_2 s_1 + M_3 s_2$$

$$\mathrm{E}[q_1|y_2] = N_0 + N_1\mathrm{E}[c_1|y_2] + N_2 s_1 + N_3 s_2.$$

Substituting Equations (C2):

$$\begin{aligned}E[q_2|y_1] &= M_0 + M_1(\beta_0 + \beta_1 c_1 + \beta_2 s_2) + M_2 s_1 + M_3 s_2 \\ &= (M_0 + M_1\beta_1) + M_1\beta_1 c_1 + M_2 s_1 + (M_1\beta_2 + \beta_3)s_2 \\ &= A_0 + A_1 c_1 + A_2 s_1 + A_3 s_2; \end{aligned} \tag{A1}$$

$$\begin{aligned}E[q_1|y_2] &= N_0 + N_1(\alpha_0 + \alpha_1 c_2 + \alpha_2 s_1) + N_2 s_1 + N_3 s_2 \\ &= (N_0 + N_1\alpha_0) + N_1\alpha_1 c_2 + (N_1\alpha_2 + N_2)s_1 + N_3 s_2 \\ &= B_0 + B_1 c_2 + B_2 s_1 + B_3 s_2. \end{aligned} \tag{A2}$$

Substituting into the expressions for q_i for $i = 1,2$ derived in the text:

$$\begin{aligned}q_1 &= (1/2)(a - c_1) - (1/2)(A_0 + A_1 c_1 + A_2 s_1 + A_3 s_2) \\ &= (1/2)(a - A_0 - (1 + A_1)c_1 - A_2 s_1 - A_3 s_2); \end{aligned} \tag{A3}$$

$$\begin{aligned}q_2 &= (1/2)(a - c_2) - (1/2)(B_0 + B_1 c_2 + B_2 s_1 + B_3 s_2) \\ &= (1/2)(a - B_0 - (1 + B_1)c_2 - B_2 s_1 - B_3 s_2). \end{aligned} \tag{A4}$$

In equilibrium, the expectation of (A3) (resp. (A4)) must coincide with our conjectures. So, taking expectations of (A3) with respect to firm 2's information set (of (A4) with respect to firm 1's information set):

$$E[q_1|y_2] = (1/2)(a - A_0 - (1 + A_1)\mu_0 - (1 + A_1)\mu_1 c_2 - [(1 + A_1)\mu_2 - A_2]s_1 - A_3 s_2)$$

$$E[q_2|y_1] = (1/2)(a - B_0 - (1 + B_1)\xi_0 - (1 + B_1)\xi_1 c_1 - B_2 s_1 - [(1 + B_1)\xi_2 + B_3]s_2 \ .$$

The proof is completed by matching the coefficients in Equations (A1) and (A3) and Equations (A2) and (A4), solving the resulting system of equations for the A's and B's and then solving for the M's and N's. ■

Proposition 2

In every linear equilibrium, (q_1^*, q_2^*) are defined as in Proposition 1 and (s_1^*, s_2^*) satisfy:

$$s_1^* = D_0 + D_1 c_1 + D_2 k_1$$

$$s_2^* = F_0 + F_1 c_2 + F_2 k_2,$$

where:

$$E[c_2|c_1] = \mu_0 + \mu_1 c_1 \qquad E[c_1|c_2] = \delta_0 + \delta_1 c_2$$

$$D_0 = \frac{1}{\varepsilon_1 - A_2^2}[-A_2(a - A_0 + A_3 F_0 + A_3 F_1 \mu_0)] \qquad F_0 = \frac{1}{\varepsilon_2 - B_3^2}[-B_3(a - B_0 + B_2 D_0 - B_2 D_1 \delta_0)]$$

$$D_1 = \frac{1}{\varepsilon_1 - A_2^2}[A_2(1 + A_1) - A_2 A_3 F_1 \mu_1 + \varepsilon_1] \qquad F_1 = \frac{1}{\varepsilon_2 - B_3^2}[B_3(1 + B_1) - B_2 B_3 D_1 \delta_1 + \varepsilon_2]$$

$$D_2 = -\left(\frac{1}{\varepsilon_1 - A_2^2}\right) \qquad F_2 = -\left(\frac{1}{\varepsilon_2 - B_3^2}\right).$$

Proof

To find a linear equilibrium, compute using the conjectured equilibrium strategies:

$$\mathrm{E}[s_2|c_1] = F_0 + F_1\mathrm{E}[c_2|c_1]$$

$$\mathrm{E}[s_1|c_2] = D_0 + D_1\mathrm{E}[c_1|c_2],$$

where $\mathrm{E}[k_j \mid k_i, c_i] = \mathrm{E}[k_j \mid c_i] = \mathrm{E}[k_j] = 0$ by independence and our assumption that $\mathrm{E}[k_i] = 0$ for $i = 1,2$. Further, since c_1, c_2 are joint normally distributed, $\mathrm{E}[c_2 \mid c_1] = \mu_0 + \mu_1 c_1$ and $\mathrm{E}[c_1 \mid c_2] = \delta_0 + \delta_1 c_2$. Substituting into Equations (4) and (5) and using the definitions of the A's and B's:

$$s_1 = \left[\frac{1}{\varepsilon_1 - A_2^2}\right][-A_2(a - A_0 + A_3F_0 + A_3F_1\mu_1) + (A_2(1 + A_1) - A_2A_3F_1\mu_1 + \varepsilon_1)c_1 - k_1]$$

$$s_2 = \left[\frac{1}{\varepsilon_2 - B_3^2}\right][-B_3(a - B_0 + B_2D_0 + B_2D_1\delta_0) + (B_3(1 + B_1) - B_2B_3D_1\delta_1 + \varepsilon_2)c_2 - k_2].$$

Since both are of the form assumed, the linear equilibrium is obtained by setting the coefficients in these equations equal to the F's and D's, respectively. ■

Proposition 3

If k_1 and k_2 are common knowledge, then there is no monotone equilibria in pure strategies.

Proof

Suppose not. Then in such an equilibrium, firm i can invert the strategy defining its rival's choice of s_j and infer c_j. Thus, the second-stage game is now a game of complete information whose solution is well-known: $q_i^* = (1/3)(a - 2c_i + c_j)$ and equilibrium profits are $(q_i^*)^2$, for $i = 1,2$. Next, consider a deviation in which firm i reports $\hat{s}_i$, which leads firm j to infer that firm i's cost of production is $\hat{c}_i \neq c_i$. Thus, $q_i = (1/3)(a - c_i - \hat{c}_i + c_j)$ and $q_j = (1/3)(a - 2c_j - \hat{c}_i)$. Since firm i's profits are q_i^2, they are decreasing in $\hat{c}_i$. As a result, if firm i deviates to $\hat{c}_i < c_i$, then its profits in the product market increase, which, for small deviations, exceed the costs of misreporting making firm i's deviation profitable and showing that there is no linear perfect Bayes equilibrium in pure strategies when k_1 and k_2 are common knowledge. ■

Proposition 4

If ε is not too small, then there is a symmetric linear perfect Bayes equilibrium in pure strategies.

Proof

In a symmetric equilibrium, $D_i = F_i$ and $\alpha_1 = \beta_i$ for $i = 1,2,3$; and $N_i = M_i$ for $i = 1,2,3,4$ and $\mu_i = \delta_i$ for $i = 1,2$. After matching coefficients, we have 12 nonlinear equations in 12 unknowns describing the coefficients of the firms' strategies. By iterative substitution, one obtains a system of two equations in two unknowns, D_1, D_2 (resp. F_1, F_2). One equation is quadratic in D_1 and has real roots when $\varepsilon_1 (= \varepsilon_2)$ is not too small. Solving for the roots and substituting into the other equation produces an 18th order polynomial in D_2 with 87 sign changes. Thus, by Descartes' Rule of Signs (see, for example, Levin [2002] and the references therein), there is at least one real root and, thus, at least one symmetric linear equilibrium in pure strategies. ■

Proof of the Comparative Static Results for Changes in ε

Suppose not. Then in equilibrium, $s_i - c_i > \xi_i > 0$ for $i = 1,2$ for all ε. Since $\mathrm{E}[\pi_i \mid k_i c_i] = \mathrm{E}[(q_i^*)^2 \mid k_i, c_i] - h_i(s_i, c_i)$ and $\mathrm{E}[(q_i^*)^2]$ is bounded from above by the firm's mo-

nopoly profits, $[(a-c_i)/2]^2$, if $s_i - c_i > \xi_i > 0$, then there is a ε large enough so that $[(a - c_i)/2]^2 - h_i(s_i, c_i) < 0$. Thus, if $s_i - c_i > \xi_i > 0$, there exists a large enough ε so that in the conjectured equilibrium, firm i's expected profits are negative. Since firm i could deviate to $s_i = c_i$ and earn non-negative profits, such a deviation is profitable and we have shown that in every linear perfect Bayes equilibrium, $s_i - c_i > \xi_i > 0$ for all ε must be false. Given this, the difference $s_i - c_i$ must become arbitrarily small as ε gets large, which implies that $E[c_i \mid c_j, s_i, s_j] - c_i$ also becomes arbitrarily small as ε becomes arbitrarily large. Finally, since q_i^* is a linear function of c_i and $E[c_j \mid c_i, s_i, s_j]$ in equilibrium, it too converges to the complete information Cournot output as ε becomes arbitrarily large. ■

REFERENCES

Arya, A., H. Frimor, and B. Mittendorf. 2010. Discretionary disclosure of proprietary information in a multi-segment firm. *Management Science* 56 (4): 645–658.

Beneish, M. 2001. Earnings management: A perspective. *Managerial Finance* 27 (12): 3–17.

Blumenstein, R., and P. Grant. 2004. On the hook: Former chief tries to redeem the calls he made at AT&T—As he retires, Armstrong blames industry fraud for breakup of company—"History Isn't Always Right." *Wall Street Journal* (May 26): A1.

Chakraborty, A., and R. Harbaugh. 2007. Comparative cheap talk. *Journal of Economic Theory* 132 (1): 70–94.

Christensen, P., and G. Feltham. 2002. *Economics of Accounting: Information in Markets*. New York, NY: Springer.

Crawford, V., and J. Sobel. 1982. Strategic information transmission. *Econometrica* 50 (6): 1431–1451.

Da Dalt, P., and S. Margetis. 2004. The intra-industry contagion effect from earnings restatements. Working paper, University of South Florida.

Darrough, M. 1993. Disclosure policy and competition: Cournot vs. Bertrand. *The Accounting Review* 68 (3): 534–561.

———, and N. Stoughton. 1990. Financial disclosure policy in an entry game. *Journal of Accounting and Economics* 12 (1–3): 219–243.

Durnev, A., and C. Mangen. 2009. Corporate investments: Learning from restatements. *Journal of Accounting Research* 47 (3): 679–720.

Feltham, G., and J. Xie. 1992. Voluntary financial disclosure in an entry game with a continua of types. *Contemporary Accounting Research* 9 (1): 46–80.

Fischer, P., and R. Verrecchia. 2000. Reporting bias. *The Accounting Review* 75 (2): 229–245.

———, and P. Stocken. 2001. Imperfect information and çredible communication. *Journal of Accounting Research* 39 (1): 119–134.

———, and ———. 2004. Effect of investor speculation on earnings management. *Journal of Accounting Research* 42 (5): 843–870.

———, and R. Verrecchia. 2004. Disclosure bias. *Journal of Accounting and Economics* 38 (3): 233–250.

Fried, D. 1984. Incentives for information production and disclosure in a duopolistic environment. *The Quarterly Journal of Economics* 99 (2): 367–381.

Fudenberg, D., and J. Tirole. 1993. *Game Theory*. Cambridge, MA: The MIT Press.

Gal-Or, E. 1986. Information transmission—Cournot and Bertrand equilibria. *The Review of Economic Studies* 53 (1): 85–92.

——— 1988. The advantages of imprecise information. *The Rand Journal of Economics* 19 (2): 266–275.

Gleason, C., N. Jenkins, and W. Johnson. 2008. The contagion effects of accounting restatements. *The Accounting Review* 83 (1): 83–110.

Gonen, I. 2003. Intra-industry effects of corrective disclosures: Is mistrust contagious? Working paper, New York University.

Guttman, I., O. Kadan, and E. Kandel. 2006. A rational expectations theory of the kink in accounting earnings. *The Accounting Review* 81 (4): 811–848.

Har, C., V. Nagar, and P. Petacchi. 2009. The effect of rational capital markets versus regulatory enforcement

on the valuation of innovative financial assets. Working paper, University of Michigan.

Kedia, S., and T. Philippon. 2009. The economics of fraudulent accounting. *Review of Financial Studies* 22 (6): 2169–2199.

Levin, S. 2002. Descartes' rule of signs—How hard can it be? Working paper, Stanford University.

Margetis, S. 2004. Straight and convertible debt and the intra-industry contagion effect of earnings restatements. Working paper, University of South Florida.

Martin, D. 2004. *Tough Calls: AT&T and the Hard Lessons Learned from the Telecom Wars*. New York, NY: AMACOM/American Management Association.

McConnell, B., J. Higgins, K. Kerschbaumer, and P. Albiniak. 2002. In the Loop. *Broadcasting and Cable Magazine* (November 25): 1–3.

Mulford, C., and E. Comiskey. 2002. *The Financial Numbers Game*. New York, NY: John Wiley & Sons.

Raith, M. 1996. A general model of information sharing in oligopoly. *Journal of Economic Theory* 71 (2): 260–288.

Ronen, J., and V. Yaari. 2008. *Earnings Management*. New York, NY: Springer Science+Business Media.

Sadka, G. 2004. Financial reporting and product markets: Learning from competitors. Working paper, University of Chicago.

——— 2006. The economic consequences of accounting fraud in product markets: Theory and a case from the U.S. telecommunications industry (WorldCom). *American Law and Economics Review* 8 (3): 439–475.

Schilit, H. 1993. *Financial Shenanigans*. Boston, MA: McGraw-Hill.

Searcy, D. 2005. Tracking the numbers/outside audit: On judgment day, assessing Ebbers's impact. *Wall Street Journal* (July 13): C1.

Shapiro, C. 1986. Exchange of information in oligopoly. *The Review of Economic Studies* 53 (3): 433–446.

Stocken, P., and R. Verrecchia. 2004. Financial reporting system choice and disclosure management. *The Accounting Review* 79 (4): 1181–1203.

Thornburgh, R. 2002. First interim report of Dick Thornburgh, bankruptcy court examiner. Available at: http://fl1.findlaw.com/news.findlaw.com/hdocs/docs/worldcom/thornburgh1strpt.pdf.

Verrecchia, R. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5 (1): 179–194.

——— 2001. Essays on disclosure. *Journal of Accounting and Economics* 32 (1–3): 97–180.

Vives, X. 2002. Private information, strategic behavior, and efficiency in Cournot markets. *The Rand Journal of Economics* 33 (3): 361–376.

——— 2006. Information sharing among firms. Working paper, University of Navarra.

Wagenhofer, A. 1990. Voluntary disclosure with a strategic opponent. *Journal of Accounting and Economics* 12 (1-3): 341–363.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory*. London, U.K.: Prentice Hall.

———, and ———. 1990. Positive accounting theory: A ten year perspective. *The Accounting Review* 65 (1): 131–156.

Xu, T., M. Najand, and D. Ziegenfuss. 2006. Intra-industry effects of earnings restatements due to accounting irregularities. *Journal of Business Finance & Accounting* 33 (5): 696–714.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1215–1238

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1215

Financial Reporting Quality, Private Information, Monitoring, and the Lease-versus-Buy Decision

Anne Beatty
The Ohio State University

Scott Liao
University of Toronto

Joseph Weber
Massachusetts Institute of Technology

ABSTRACT: A flourishing stream of research suggests that liquidity-constrained firms with low accounting quality have limited access to capital for investments. We extend this research by investigating whether these firms are more likely to lease their assets. Lessors' superior control rights allow them to provide capital to constrained firms with low-quality accounting reports. Consistent with this conjecture, we find that low accounting quality firms have a higher propensity to lease than purchase assets. To verify that leasing does not merely reflect these firms' desire for off-balance-sheet accounting, we investigate whether banks' access to private information and monitoring affect the relation between accounting quality and leasing. We find the association between accounting quality and leasing decreases when banks have higher monitoring incentives and when loans contain capital expenditure provisions. These results suggest that other mechanisms can substitute for the role of accounting quality in reducing information problems.

Keywords: *financial reporting quality; private information; monitoring; lease-versus-buy.*

Data Availability: *Data are available from sources identified in the text.*

We thank Wayne Thomas and two anonymous reviewers for many helpful suggestions that have resulted in substantial improvements to the manuscript.

Editor's note: Accepted by Wayne Thomas.

Submitted: August 2008
Accepted: November 2009
Published Online: June 2010

I. INTRODUCTION

A growing body of accounting research concludes that higher financial reporting quality reduces the negative effects of financing constraints on investment by mitigating information asymmetry (e.g., Biddle et al. 2009; Lara et al. 2009; Biddle and Hilary 2006; Verdi 2006; Bushman et al. 2008; Hope et al. 2009).[1] The research examining whether accounting quality reduces financing frictions and underinvestment focuses on asset purchases and R&D without considering alternative investment mechanisms, like leases. In another stream of literature, both Sharpe and Nguyen (1995) and Eisfeldt and Rampini (2009) argue that leases, which account for about one-third of new equipment investment for U.S. firms, provide creditors with more security, higher priority in bankruptcy, and an effective way of reducing adverse selection and moral hazard problems that arise from information asymmetries. Although these studies suggest that firms facing greater financing constraints have a higher propensity to engage in off-balance-sheet lease investments, they do not consider other mechanisms that reduce financing constraints, like accounting quality. We extend these two research streams by considering the role of accounting quality in the firm's propensity to lease assets (i.e., the proportion of assets that the firm leases).

We focus on the determinants of firms' propensities to acquire assets using operating leases (measured by the ratio of the capitalized minimum operating lease payment to the sum of that amount and the amount of property, plant, and equipment [PP&E]). By capitalizing operating leases, this measure is unlikely to be affected by the differences in accounting treatment for leased versus purchased assets. We conduct our analyses using a comprehensive sample of 3,033 manufacturing firms on the Compustat database. Our focus on manufacturing firms is consistent with previous research examining how financial constraints affect asset purchases. In addition, focusing on a single industry reduces concerns that our results are driven by industry-induced correlations between leasing, financing, and accounting.[2] To measure our primary variable of interest, financial reporting quality, we follow Bharath et al. (2008) by using the first principal component of three accrual-based accounting quality metrics.[3] We conduct cross-sectional regressions, examining the relationship between leasing propensity and accounting quality, controlling for size, tax incentives, financing constraints, and other factors expected to affect leasing decisions. We average all variables across the sample period, thereby treating each firm as one observation rather than using panel data.[4]

We find that firms with worse financial reporting quality have a greater propensity to lease their assets. This result highlights the importance of considering leasing when studying the effects of financial reporting quality on investments. Prior research concludes that firms with poor accounting quality have financial constraints, inhibiting their ability to invest in the periods in which they are liquidity-constrained. Our results extend this line of research, suggesting that although low-quality accounting firms are likely to forgo purchasing assets, they are more likely to lease assets in the periods in which they are liquidity-constrained.

One interpretation of our finding that poor accounting quality firms lease more of their assets is that these firms substitute lease financing for other forms of secured borrowing. An alternative

---

[1] See Hope et al. (2009) for a discussion of whether accounting quality (transparency) affects firms' ability to gain access to financing.

[2] The statistical significances of the coefficients on our variables of interest are not sensitive to this choice.

[3] As discussed in the Appendix, the three measures are Teoh et al.'s (1998) absolute value of current discretionary accruals, Dechow et al.'s (1995) absolute value of total discretionary accruals, and the cross-sectional version of Dechow and Dichev's (2002) absolute value of current discretionary accruals.

[4] Difficulties measuring inter-temporal variation in accounting quality and lease propensity make cross-sectional regressions more appropriate than time-series regressions in our setting. Biddle et al. (2009) use a panel data set to investigate the relationship between investment, the business cycle, and accounting quality.

explanation of our finding is that the negative association between the proportion of assets leased and accounting quality merely reflects the desire of low accounting quality firms to engage in off-balance-sheet financing. To distinguish between these alternatives, we investigate how other mechanisms that reduce information asymmetry and agency problems impact the relationship between accounting quality and leasing propensity. Specifically, we investigate how lenders' monitoring and loan covenants affect the association between accounting quality and leasing propensity. If the association between accounting quality and leasing reflects the role of accounting quality in reducing information asymmetry, then we would expect the negative relation between accounting quality and leasing to be mitigated by other information asymmetry and agency-cost-reducing mechanisms, suggesting that accounting quality and leasing are alternative ways to address financial constraints. Alternatively, if off-balance-sheet accounting motivates firms with poor accounting quality to lease, then the relationship between leasing and accounting quality should not be affected by other information-asymmetry-reducing mechanisms.

This analysis also allows us to extend Biddle and Hilary's (2006) investigation of whether cross-sectional differences in access and use of private information and monitoring affect the role of accounting quality in investment decisions. Specifically, Biddle and Hilary (2006) suggest that the role of accounting quality in capital investment decisions depends on the firm's use of other monitoring mechanisms that reduce information asymmetry. We expand this reasoning, suggesting that both the quality of banks' monitoring and the covenants banks use to restrict investment will affect the importance of accounting quality in the lease-versus-buy decision.

To test these hypotheses, we use a subsample of manufacturing firms with syndicated bank loan data on either the Loan Pricing Corporation (LPC) or Security Data Corporation (SDC) databases. There are two primary reasons why we focus on this important subsample of the debt market.[5] First, the variation in bank loan ownership provides cross-sectional variation in lenders' monitoring incentives. Lead lenders with proper financial incentives to conduct due diligence reviews prior to loan initiation and to provide loan monitoring over the life of the contract have incentives to reduce the agency problems induced by information asymmetries between lenders and opaque borrowers. Therefore, we hypothesize that when lead lenders have greater due diligence and monitoring incentives, the importance of the quality of accounting information in the lease-versus-buy decision will be mitigated.

The second advantage of focusing on syndicated loans is that, unlike public debt, there is cross-sectional variation in the use of contractual limitations on asset purchases in these loans.[6] When a firm's debt contract contains an investment-restriction covenant designed to mitigate the agency problems of debt, we expect the quality of accounting information will be less important in controlling information problems. Accordingly, we hypothesize that the effect of accounting quality in the lease-versus-buy decision decreases in the face of investment-restriction covenants.

We identify a subsample of 997 manufacturing firms with loan ownership data in the LPC database to test our predictions about the effects of banks' direct and indirect monitoring on the role of accounting quality in the lease-versus-buy decision. Based on Sufi (2007) and Ball et al. (2008), we use lead lender ownership to proxy for lenders' due diligence and monitoring incentives.[7] Consistent with bank monitoring substituting for accounting information, we find that accounting quality is less important in the lease-versus-buy decision when lead lenders' ownership is high.

---

[5] Sufi (2007) reports that syndicated loans represent 51 percent of U.S. corporate finance originations, and generate more underwriting fees than both equity and debt underwriting. He also notes that almost 90 percent of the largest 500 nonfinancial Compustat firms obtained a syndicated loan between 1994 and 2002.

[6] Nini et al. (2009) find that 32 percent of U.S. bank loans contain this covenant, while we find that less than 5 percent of public debt uses this covenant to directly restrict firms' opportunistic investments.

[7] In sensitivity tests, Sufi (2007) documents that the lead lender ownership measure is very robust.

We then examine the effects of contractual restrictions on asset purchases, using a subsample of 865 manufacturing firms with information on the investment-restriction covenant in the SDC database. We use an indicator variable to capture the effects of covenants on leasing decisions. Consistent with our previous analysis, we find that the magnitude of the correlation between our accounting quality measure and proportion of assets leased is lower in the presence of this covenant, suggesting that this covenant diminishes the importance of accounting information in the lease-versus-buy decision.

We examine the sensitivity of our results to a variety of different research design choices. Specifically, we consider the effect of limiting the sample to manufacturing firms and the choice of measures for leasing propensity, accounting quality, and lead lender measures. We also investigate the sensitivity of our results to the methods we use to address endogeneity and the methods we use to address missing lease data. Finally, we investigate whether our results are affected when we use an alternative research design to address cross-sectional variation in financing constraints. Each of these research design modifications produces robust results.

Our study contributes to the accounting literature examining the role of accounting quality in investment decisions, and to the finance literature examining the determinants of the decision to lease-versus-buy assets, by showing that firms with information asymmetries due to poor accounting quality are more likely to lease their assets. Results also suggest that accounting quality is important in mitigating the financing constraints that lead firms to lease rather than buy assets.

Our setting also allows us to test the conjecture made in previous research that banks' use of covenants, access to private information, and monitoring are substitutes for accounting information in reducing information asymmetries. Results suggest that bank monitoring and covenant restrictions are substitutes for the influence of accounting quality on investment decisions.

Section II provides the background to our study, Section III contains the hypothesis development, Section IV discusses the sample selection and database, Section V displays the model specification, Section VI presents the results, Section VII reports robustness checks, and Section VIII concludes.

II. BACKGROUND

Information Asymmetry and Investment

Fazzari et al. (1988) examine the importance of capital market imperfections caused by information asymmetry between managers and outside capital providers on firms' optimal investment decisions. They find that financially constrained firms rely on internally generated cash flows to fund their investments, increasing the sensitivity of their investments to cash flows.[8]

Many studies, such as Bushman et al. (2005), Biddle and Hilary (2006), Verdi (2006), Biddle et al. (2009), and Lara et al. (2009), suggest that aspects of the firms' accounting reports may affect the extent to which firms' information asymmetry problems affect their investments. Despite the fact that these studies have used a variety of measures of both investment efficiency and accounting quality, the results of these studies have all been consistent. When accounting quality is high, the reduction in information asymmetry leads to an improvement in investments.

Bushman et al. (2005) investigate cross-country variation in the timeliness of loss recognition and the timeliness of change in investment policy. They find that firms in countries with more timely loss recognition respond more quickly to declining investment opportunities by reducing net inflows of capital to new investments. Their study implies that better accounting quality leads to increased investment efficiency.

[8] Although many studies conclude that information asymmetry and financing constraints decrease investment efficiency, using investment-cash flow sensitivity to proxy for investment efficiency has been controversial (e.g., Kaplan and Zingales 1997, 2000; Bushman et al. 2008).

Biddle and Hilary (2006) examine how accounting quality affects firms' investment-cash flow sensitivity.[9] They examine the effect of accounting quality on investment shocks, finding that higher accounting quality is associated with lower investment-cash flow sensitivity. They also find that accounting quality plays a relatively more important role in economies where the capital is primarily provided through arm's-length transactions and investors do not have access to private information. At the extreme, Biddle and Hilary (2006) find that in Japan, where capital is typically provided by keiretsu and banks, accounting quality does not affect the association between investment and cash flows. One of the implications of Biddle and Hilary's (2006) results is that firms with poor accounting quality are more likely to rely on internally generated cash flows to fund their investments, and thus have more difficulty obtaining external financing to fund their investments. Their results also imply that the role of accounting quality in capital investment decisions depends on whether the firm uses other monitoring mechanisms that reduce information asymmetry.

Verdi (2006) and Biddle et al. (2009) are concerned with both whether firms overinvest in losing projects and whether they underinvest in positive net-present-value projects. Both Verdi (2006) and Biddle et al. (2009) find that financial reporting quality is positively correlated with investment when firms are prone to underinvest and is negatively correlated with investment when firms are prone to overinvest. Biddle et al. (2009) conclude that, "higher reporting quality is associated with both lower over- and underinvestment." Lara et al. (2009), draw similar conclusions from a related study using an alternative measure of accounting quality.[10]

These studies focus on the role of accounting quality on firms' decisions to purchase assets (like PP&E). A related literature examines the role of accounting quality on firms' ability to obtain traditional financing (like debt or equity).[11] These studies find that firms with poor accounting quality have difficulty raising capital from either debt or equity markets, and thus invest suboptimally. For example, Bharath et al. (2008) find that firms with poor accounting quality have to pay higher interest rates. Further, Francis et al. (2004) find that firms with better accounting quality bear lower cost of equity.[12] Most recently, Hope et al. (2009) investigate the relationship between financial reporting transparency and firms' ability to access capital. They identify a sample of privately held firms with cross-sectional variation in whether an audit has been conducted, and find that audited firms with more transparent financial statements have better access to external financing. Their results imply that financial transparency influences information asymmetries, and thus influence firms' access to capital from traditional debt markets. Thus, firms with lower quality accounting are inefficient (have lower investment efficiency). We extend this research by investigating whether firms with lower quality accounting are precluded from investment, or whether they can use an alternative form of financing (an operating lease).

Leasing and Financing Constraints

Eisfeldt and Rampini (2009) and Sharpe and Nguyen (1995) find that firms with larger information asymmetries (greater financing constraints) are more likely to obtain their assets through leases. Sharpe and Nguyen (1995) suggest that institutional features specific to operating

---

[9] Biddle and Hilary (2006) employ Dechow and Dichev's (2002) measure to capture accounting quality.

[10] Verdi (2006) and Biddle et al. (2009) augment the Dechow and Dichev (2002) measure of accounting quality according to McNichols (2002). Lara et al. (2009) use conditional conservatism to capture accounting quality.

[11] The examination of the effects of financing constraints on investment in balance sheet assets (like PP&E) is not unique to these studies. Most of the related finance research follows this same approach.

[12] Francis et al. (2004) use multiple variables to measure accounting quality, including earnings persistence, predictability, smoothness, relevance, timeliness, accounting conservatism, and accrual quality measured using the Dechow and Dichev (2002) model.

leases provide lessors with important advantages in bankruptcy.[13] If the lessee goes bankrupt, then the lessee can either accept or breach all lease obligations. If the lessee accepts the obligations, then the lessor is entitled to continue receiving payments based on the original lease agreement. The rental expenses, along with lessors' other claims, are classified as administrative expenses in the bankruptcy code, which must be paid immediately or "within a reasonable period."[14] In this case, the bankruptcy filing does not change the status of the lessor's claims. Alternatively, if the lessee breaches the lease obligations, then the lessor can immediately repossess the property. Any economic losses accompanied with this repossession, including unpaid rents, late fees, or the loss regarding the re-lease or sale of the property, is an unsecured claim against the lessee. In contrast, other creditors are left with no assurance that their claims will be recovered.[15]

Sharpe and Nguyen (1995) argue that the preferential treatment of leases in bankruptcy causes financially constrained firms to prefer to lease assets rather than borrow to buy assets. Consistent with this prediction, they find firms that are non-dividend-paying, lower-rated, or cash-poor are more inclined to lease than nonfinancially constrained firms. This suggests that firms with information asymmetries due to poor accounting quality may prefer other alternatives, like leases.

Banks' Private Information and Monitoring

Biddle and Hilary (2006) suggest that there are other mechanisms that may substitute for accounting quality in reducing information asymmetry, thereby alleviating financing restrictions and increasing investment efficiency. In contrast to their findings for the U.S., Biddle and Hilary (2006) find that investment-cash flow sensitivities do not vary with accounting quality in Japan. They conjecture that, unlike U.S. markets, Japanese capital markets resolve information asymmetry problems through relationship banking, such that public information loses its importance in affecting firm operations. Biddle and Hilary (2006) argue that banks' superior private information should serve as a substitute for accounting quality in determining the sensitivity of investment to internally generated cash flows.[16]

In addition to having access to private information, Nini et al. (2009) suggest that banks often include loan covenants setting an upper limit on capital expenditures and find that this capital expenditure restriction provision successfully reduces firms' capital expenditures. They argue that the investment-restriction covenant mitigates the asset substitution and overinvestment problems. Typically, these covenants define capital expenditures as assets included in the property, plant, and equipment accounts under GAAP. Therefore, the covenants would restrict asset purchases and capital leases, but not operating leases, and are thus likely to affect both the use of operating leases and the importance of accounting information in the leasing decision.

The following excerpt from Georgia Pacific's December 2005 syndicated loan agreement is an illustration of a typical capital expenditure restriction:

---

[13] SFAS No. 13 requires leases to be classified as either operating or capital. A lease meeting any of the following four criteria is classified as a capital lease: (1) ownership of the leased asset is transferred to the lessee at lease maturity, (2) a bargain purchase option is available, (3) the lease term is equal to 75 percent or more of the remaining economic life of the leased asset, and (4) the present value of the minimum lease payments equals or exceeds 90 percent of the asset's market value.

[14] Lessors' other claims might include delinquencies, late fees, and other damages suffered.

[15] This provision applies to operating leases, but not capital leases. If the original lease contract stipulates that the lessor does not retain property ownership at lease maturity, then the lease is treated as secured debt.

[16] This prediction is consistent with theoretical models such as Holthausen and Verrecchia (1988).

FINANCIAL COVENANTS

Maximum Capital Expenditures of up to:

Fiscal Year	Capital Expenditures
2006	$900,000,000
2007	$1,100,000,000
2008	$1,100,000,000
2009	$900,000,000
2010 and thereafter	$700,000,000

The lesser of (a) 100 percent of the unused amounts allocated to any year and (b) 50 percent of the maximum Capital Expenditures permitted in such year shall be available for Capital Expenditures in the following fiscal year only. In addition, the Capital Expenditure covenant will permit the amount permitted, but not used, for Restricted Payments and investments in joint ventures under clause (12) of the definition of "Permitted Investments" to be used to make additional Capital Expenditures. Capital Expenditures shall exclude, whether such a designation would be in conformity with GAAP, expenditures made in connection with the replacement or restoration of property, to the extent reimbursed or financed from insurance or condemnation proceeds.

CERTAIN DEFINITIONS

"***Capital Expenditures***" means, for any period and with respect to any Person, any and all expenditures made by the Parent or any of its Subsidiaries in such period for assets added to or reflected in its property, plant, and equipment accounts or other similar capital asset accounts or comparable items or any other capital expenditures that are, or should be, set forth as "additions to plant, property, and equipment" on the consolidated financial statements of the Parent prepared in accordance with GAAP, whether such asset is purchased for cash or financed as an account payable or by the incurrence of Indebtedness, accrued as a liability or otherwise.

III. HYPOTHESIS DEVELOPMENT

Two streams of research consider how firms mitigate the effects of financing constraints on investment. The first suggests that firms can reduce information asymmetry and financing constraints by committing to better accounting quality. The second finds that firms with larger financing constraints lease more of their assets. Together, these two streams of research lead us to predict that firms with worse accounting quality are more likely to obtain their assets through operating leases.

H1: A firm's propensity to lease-versus-buy assets is declining in its accounting quality.

An alternative explanation for a negative association between a firm's accounting quality and its propensity to enter into operating leases is that firms with poor accounting quality prefer off-balance-sheet financing. Under this alternative, we would expect the relationship between leasing and accounting quality to be unaffected by other mechanisms that reduce information asymmetry.

One mechanism that is likely to affect the extent of the firm's information asymmetry problem is whether the outside suppliers of capital, such as banks, have incentives to obtain information through private channels and thus reduce information asymmetry. If better accounting quality reduces information asymmetry and thus reduces a firm's leasing propensity, then other mecha-

nisms that reduce information asymmetry like the lenders' incentives to access private information and monitor the borrower are likely to affect the importance of accounting quality on the lease-versus-buy decision, motivating our second hypothesis.

H2: The importance of accounting quality in the lease-versus-buy decision decreases with banks' incentives to perform due diligence.

Capital expenditure covenants are another mechanism that lenders use to reduce agency costs.[17] If these covenants are effective at reducing agency costs and financing constraints, then the importance of accounting quality in the lease-versus-buy decision will also be reduced by capital expenditure covenants. In other words, accounting quality affects the lease-versus-buy decision (i.e., low-quality accounting limits or makes it more costly for firms to engage in capital spending so they turn to leasing). However, capital expenditure covenants, by their very nature, also limit the ability of firms to engage in capital spending. Therefore, when a covenant is already in place, the incremental impact of accounting quality on the lease-versus-buy decision is reduced. For example, a firm with a covenant in place could have to opt for leasing, even though it has high-quality accounting.[18]

Similar to our monitoring hypothesis, the negative association between a firm's accounting quality and their propensity to enter into operating leases could be due to firms with poor accounting quality preferring off-balance-sheet financing. Under this alternative hypothesis, we would expect the relationship between leasing and accounting quality to be unaffected by the use of capital investment restrictions. We test this reasoning in our third hypothesis.

H3: The importance of accounting quality in the lease-versus-buy decision is lower in the presence of capital expenditure restriction provisions.

IV. SAMPLE SELECTION

We use three different samples to test our hypotheses. The first is a comprehensive sample of manufacturing firms included on Compustat. Our second and third samples are subsamples of the first sample. The first subsample is restricted to firms with data on lead lenders contained on the LPC database. The second subsample requires data on capital expenditure covenants on the SDC database.

Comprehensive Sample

We test our first hypothesis on the role of accounting quality in the leasing decision using a comprehensive sample of Compustat manufacturing firms, identifying 3,033 firms in SIC codes (2000–3999) during the period 1995–2006 with the necessary data.

Subsamples with Lead Lender and Capital Expenditure Covenant Data

We use a subsample of firms that borrow in the syndicated loan market to probe the cross-sectional variation in the effect of accounting quality on leasing activities and to ensure that our results are not an artifact of a desire for off-balance-sheet accounting. We focus on the syndicated loan market for two reasons. First, the variation in bank loan ownership provides cross-sectional variation in lenders' private information and monitoring incentives. As we discuss above, the

[17] We focus on this covenant rather than other financial covenants for two reasons. First, this covenant restricting asset purchases impacts the lease-versus-buy decision on which we focus. Second, this covenant is not based on any financial numbers, so it is unaffected by the method of accounting for operating leases.

[18] This argument might suggest that this covenant should not be required for high-accounting quality firms. In our investment covenant model (see Table 7) we find that beyond accounting quality there are many other important determinants of these covenants suggesting high accounting quality firms might have these covenants.

syndicate structure provides loan owners varying incentives to incur costs to collect private information and to monitor management (Sufi 2007). Therefore, lead arrangers' ownership is an established proxy for the private information acquired and the quality of monitoring imposed by banks.

The second reason we focus on this market rests on the use of debt covenants in this market. Since private lenders have informational and monitoring advantage over public debtholders in writing and enforcing debt covenants that mitigate the agency problems of debt (e.g., Fama 1985; Diamond 1984), bank loans provide cross-sectional variation in the use of debt covenants that is rare in corporate bonds. We exploit the variation of the use of the covenant restricting capital expenditures to investigate the effect of direct monitoring by banks on the lease-versus-buy decision.

To test the relation between accounting quality and banks' private information and monitoring quality, we use a sample of 997 manufacturing firms that are covered by the LPC database and have loan ownership data. Consistent with prior research, to test our third hypothesis we use a sample of firms listed in the SDC syndicated loan database to identify firms that have investment-restriction covenants. We identified 865 firms that are listed in the SDC database.

V. RESEARCH DESIGN AND PROXIES

The basic model that we use to test all three of our hypotheses is the same, although the samples studied and the estimation techniques used to estimate the models differ across our tests. To test our hypotheses, we estimate the following model:

$$Lease = \beta_0 + \beta_1 * AccQual + \beta_2 * Nodiv + \beta_3 * Unrated + \beta_4 * Rating + \beta_5 * Size + \beta_6 * Loss + \beta_7 * MTR + \varepsilon \quad (1)$$

where:[19]

Lease = the average of the capitalized lease expenditure (lagged Compustat #96 multiplied by 10) over the sum of PP&E and capitalized lease expenditure (Compustat #8 + lagged Compustat #96 × 10);

AccQual = the first principal component of three accrual-based accounting quality metrics based on Bharath et al. (2008; described in the Appendix);

Nodiv = an indicator variable equal to 1 if the firm never pays dividends (Compustat #21) during the sample period, and 0 otherwise;

Unrated = an indicator variable equal to 1 if the firm is never rated by the S&P (Compustat #280) during the sample period, and 0 otherwise;

Rating = the average of S&P ratings when the firm is rated by the S&P; following Sharpe and Nguyen (1995), the rating is coded 1 if the S&P rating is between AAA and A–, 2 if the rating is between A+ and A−, 3 if the rating is between BBB+ and BBB–, and 4 if the rating is between BB+ and D;

Size = the decile ranking of average of the natural log of sales (Compustat #12);

Loss = the average of the indicator variable for loss year, where the indicator variable equals 1 when the Compustat #123 is less than 0; and

MTR = the average of the marginal tax rates, based on John Graham's simulation technique, as reported on his website.[20]

[19] All OLS t-statistics are based on standard errors clustered at the three-digit industry level.

[20] To see a summary of the approach he uses to obtain marginal tax rates, and request access to the underlying data, please visit the following website: http://faculty.fuqua.duke.edu/~jgraham/.

We capitalize the minimum lease payment assuming a cost of capital of 10 percent and that the lease payments are a perpetuity.[21] As such, the measure will be unaffected by the differences in accounting treatment for these two investment strategies.[22]

Accounting quality has numerous dimensions (Dechow et al. 2009). The literature arguing that accounting quality reduces investment frictions focuses on accounting quality's ability to reduce information asymmetries (Biddle et al. 2009). This literature is not concerned with whether accounting quality reflects lower earnings management, less business uncertainty, more stable cash flows, or some other factor that reduces information asymmetry. Similar to the prior research, we do not distinguish among these sources of accounting quality. Instead we follow Bharath et al. (2008), using a principal-component analysis to isolate the common component of three common firm level measures of accounting quality (*AccQual*). The accruals measures are based on Teoh et al. (1998), Dechow et al. (1995), and Dechow and Dichev (2002). The Appendix includes a description of our implementation of the Bharath et al. (2008) measure.

Our definition of variables differs from theirs slightly because we modify the deflator used in constructing these variables to accommodate differences in accounting for operating leases versus asset purchase. The total asset deflator used in prior research would be affected by the off-balance-sheet accounting treatment for operating leases. Instead, we use the sum of PP&E and capitalized minimum lease payment, which is the same deflator used in calculating our dependent variable (see Christie 1987).[23]

The lease-versus-buy decision is likely to be affected by a variety of other factors (see Smith and Wakeman 1985). Taxes are one of the primary motivations for leasing. Firms that are not paying taxes (or have low marginal tax rates) prefer to lease as opposed to buy, as leasing allows firms in a higher tax bracket to take more advantage of the depreciation deduction.

Leasing is also likely to be related to whether the firm faces financing constraints. Following Sharpe and Nguyen (1995) and Eisfeldt and Rampini (2009), we include three variables to capture the financing constraints: *Nodiv*, *Unrated*, and *Rating*. Non-dividend-paying firms are expected to have more financing constraints and, therefore, are more likely to lease. Unrated firms are also expected to have higher leasing propensity. We expect that worse-rated firms have greater difficulties in raising capital, so this variable should be positively correlated with *Lease*.

We also include *Size* and *Loss* as control variables. Larger firms are likely to be less financially constrained and, therefore, *Size* is expected to have a negative coefficient. Finally, we assume loss firms have a lower marginal tax rate and are thus unable to take full advantage of the tax benefits of ownership of equipment. Therefore, we expect *Loss* to be positively correlated with lease activity.

It is important to note that we average all variables across the sample period, thereby treating each firm as one observation rather than using panel data. There are several justifications for this research design choice. First, our primary variable of interest is a stock measure of the leasing propensity. This measure, as well as our financial reporting quality measure, is highly correlated through time. Averaging our variables reduces the independence problems that arise from the

[21] While parsimonious, we recognize that this approach has drawbacks because the exact discount rate and payment timing is not known. To address these concerns, we also construct the variable discounting all future lease payments disclosed in the lease footnotes (five years of payments and the thereafter piece), and varying the discount rate (between 8 and 12 percent) and the results of our main tests are consistent with those reported. Thus, our capitalization assumptions do not appear to be driving our results.

[22] In our study, leases represent operating leases and therefore use of Compustat #96 to measure lease activities should be appropriate. Since we classify capital leases as borrow-and-buy and since Compustat #8 includes capital lease expenditures, our choice of this variable should be comprehensive.

[23] We also use total assets, total assets excluding property, plant, and equipment (Compustat #8) and sales (Compustat #12) as deflators, the results in the main analysis continue to hold.

serial correlation. Although this approach only captures the between-firms effect, it eliminates within-firm correlation issues. We investigate the sensitivity of our results to this research design choice using a variety of analyses that are discussed in depth in the sensitivity analysis section. In particular, we replace the stock measure of leasing with a flow measure, and replace our accounting quality measure with a measure using three years of data prior to the leasing decision, finding similar results.

Comprehensive Sample

To test H1, we estimate Equation (1) using OLS estimation for our comprehensive Compustat sample of manufacturing firms averaged over 1995–2006. If accounting quality and leasing are two alternative mechanisms for dealing with financing constraints, then we expect better accounting quality to be associated with a lower propensity to lease (i.e., β_1 should be negative).

Analysis of Due Diligence Efforts

To test H2, we follow Bharath et al. (2008) and estimate an endogenous switching model (Lee 1978) for our subsample of 997 manufacturing firms that are covered by the LPC database and have loan ownership data. The endogenous switching model allows us to control for the endogeneity of the syndicate participants' choice to require a larger lead arranger ownership position in the loan to enhance their due diligence and monitoring efforts when the borrower has more serious information asymmetries and higher credit risk (Sufi 2007).

More specifically, the endogenous switching model has a "selection" equation modeling the syndication's decision to require lead lenders to hold a significant ownership position and two second-stage "regime" equations modeling the determinants of the firm's leasing propensity (one for firms where the lead lender holds a significant position, and another for firms where the lead lender holds a relatively small amount of the loan). The three regressions are estimated simultaneously using full information maximum likelihood. As discussed in Bharath et al. (2008), this approach controls for the self-selection associated with the decision by the syndicate to require the lead arranger to hold a relatively larger proportion of the loan and allows us to compare the effects of accounting quality on the firm's leasing propensity across the two "regimes." It also has the benefit of not requiring the coefficients on the determinants to lease to be the same across the two different regimes.

The first-stage probit model of the proportion of the loan held by the lead bank is as follows:

$$\begin{aligned} Lead\ Rank = {} & \beta_0 + \beta_1 * AccQual + \beta_2 * Nodiv + \beta_3 * Unrated + \beta_4 * Rating + \beta_5 * Size \\ & + \beta_6 * Loss + \beta_7 * MTR + \beta_8 * Loan + \beta_9 * Maturity + \varepsilon. \end{aligned} \quad (2)$$

Most of the variables in Equation (2) are also included in Equation (1) and are defined above. To capture lead banks' incentives to acquire private information and monitor management we develop the measure *Lead Rank*. This variable is an indicator variable equal to 1 if the average of the lead arrangers' ownership of the firm's bank loans exceeds the median of all sample firms, and 0 otherwise.[24] We also include two exogenous variables (exogenous in the sense that they are likely to be associated with the syndicate's decision to require the lead lender to hold a significant proportion of the loan, but not associated with the firm's leasing propensity). The first exogenous variable is the size of the deal (*Loan*) and the second exogenous variable is the maturity of the debt

[24] Since firms may have more than one loan during the sample period, lead arranger ownership and other loan characteristics are weighted (by deal amount) and averaged over all bank loans issued during the sample period. Using equally weighted values does not change the tenor of the results. We examine the sensitivity of our results to the median cut-off in the robustness section.

(*Maturity*). The choice of these two instruments is motivated by Sufi (2007). Larger loans and loans with longer maturities are likely to be made to borrowers with better credit quality, reducing the importance of the monitoring provided by the lead lender. This suggests a negative relationship between loan size, maturity, and the extent of the lead lender's holdings. Alternatively, larger loans and loans with longer maturities are likely to require more monitoring, suggesting that syndicate members might require the lead lender to hold larger shares of these loans, suggesting a positive relation. Thus, we do not predict a signed relationship between the lead lender's share of the loan, maturity, and the size of the loan.

We use maximum likelihood estimation for our two "regime" regressions. If debtholder monitoring is a potential substitute for accounting quality in solving information problems, then, in the regime regressions, we expect the effect of accounting quality (*AccQual*) on the leasing propensity to be relatively larger for firms where the lead lender holds a relatively smaller proportion of the loan. We test this using a Chi-squared test of the equality of the coefficient on accounting quality across the two regime regressions.

Analysis of Capital Expenditure Covenant

Similar to our tests of H2, we again employ an endogenous switching model to test H3, which predicts that the effects of investment restrictions on the importance of accounting quality on the firm's leasing propensity. The selection model we employ in this analysis is as follows:

$$Inv\_Cov = \beta_0 + \beta_1 * AccQual + \beta_2 * Nodiv + \beta_3 * Unrated + \beta_4 * Rating + \beta_5 * Size + \beta_6 * Loss + \beta_7 * MTR + \beta_8 * Loan + \beta_9 * Maturity + \varepsilon. \quad (3)$$

Most of the variables in Equation (3) are also included in Equation (1) and are defined above. *Inv_Cov* is an indicator variable that equals 1 if the firm ever includes a capital expenditure restriction provision in its bank loans during the sample period. The two "regime" regressions are the same as Equation (1), except that the partition is based on whether the firm's debt has an investment restriction. If this covenant is useful in reducing the asset substitution or overinvestment problem, then we expect the effect of accounting quality (*AccQual*) on the lease-versus-buy decision to be reduced in the presence of this restrictive covenant.

VI. RESULTS

Descriptive Statistics

Table 1 provides descriptive statistics for our comprehensive sample partitioned by whether firms are relatively more likely to lease assets (the high lease group) or less likely to lease assets (the low lease group). The partition is based on the median lease propensity in the sample. By construction, the high lease group has a higher leasing propensity than the low lease group. The high lease group also has lower accounting quality than the low lease group. They are also less likely to pay dividends, less likely to be rated, and if they are rated, then they are likely to have a poor credit rating, suggesting that firms that lease are more likely to face financing constraints. Finally, firms with a higher propensity to lease have more losses, and have lower marginal tax rates.

Pearson correlations are reported in Table 2. The variables in this table are highly correlated, in a statistical sense, with p-values less than 1 percent. Consistent with H1, *AccQual* and *Lease* are negatively correlated. Also consistent with prior discussions, firms with financing constraints are more likely to lease than to purchase assets. In addition, *Size* and most variables (*Nodiv*, *Unrated*, *Rating*, and *Loss*) are highly correlated (over 50 percent). *Loss* and *MTR* are also highly correlated.

Table 3 reports descriptive statistics for the subsample of firms with syndicated loan data. In the second and third columns of Table 3, we partition the sample firms by the level of lead

TABLE 1

Descriptive Statistics (Means and Medians) for Sample Firms Partitioned by Their Leasing Propensity for the Period 1995–2006

	High Lease		Low Lease	
	Mean	Median	Mean (t-stat for High Lease-Low Lease)	Median (Wilcoxon z-stat for High Lease-Low Lease)
Lease	0.650	0.652	0.223 (94.54)***	0.223 (47.69)***
AccQual	−0.147	0.079	0.217 (−18.11)***	0.348 (−22.63)***
Nodiv	0.822	1.000	0.464 (22.15)***	0.000 (20.55)***
Unrated	0.880	1.000	0.584 (19.50)***	1.000 (18.38)***
Rating[a]	3.604	4.000	3.134 (6.53)***	3.250 (6.67)***
Size	3.194	3.000	5.807 (−28.14)***	6.000 (−25.06)***
Loss	0.580	0.600	0.282 (25.73)***	0.200 (23.23)***
MTR	0.108	0.054	0.198 (−20.58)***	0.214 (−20.20)***
n	1,517	1,517	1,516	1,516

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively. High Lease (Low Lease) include firms with Lease above (below) the median.

[a] The mean (median) values of this variable only apply to firms rated by the S&P.

Variable Definitions:

Lease = average of the capitalized lease expenditure (lagged Compustat #96 × 10) over the sum of PP&E and capitalized lease expenditure (Compustat #8 + lagged Compustat #96 × 10);

AccQual = average of the first principal component of three accrual-based accounting quality metrics based on Bharath et al. (2008, described in the Appendix);

Nodiv = indicator variable equal to 1 if the firm never pays dividend (Compustat #21) during the sample period, and 0 otherwise;

Unrated = indicator variable equal to 1 if the firm is never rated by the S&P (Compustat #280) during the sample period, and 0 otherwise;

Rating = average of S&P ratings when the firm is rated by the S&P. Following Sharpe and Nguyen (1995), the rating is coded 1 if the S&P rating is between AAA and A−, 2 if the rating is between A+ and A−, 3 if the rating is between BBB+ and BBB−, and 4 if the rating is between BB+ and D;

Size = decile ranking of average of the natural log of sales (Compustat #12);

Loss = average of the indicator variable for loss year, where the indicator variable equals 1 when the Compustat #123 is less than 0; and

MTR = average of the marginal tax rates, based on John Graham's simulation technique, as reported on his website.

arranger ownership for the LPC subsample. Firms with better private information and monitoring quality tend to be smaller, have worse accounting quality, and face more financial constraints. These firms are also characterized by a smaller debt size and shorter maturities. The fact that these firms have a greater propensity to lease is consistent with Sufi's (2007) finding that a high lead arranger ownership is associated with high information asymmetries and financing constraints. This observation suggests that it is important to control for endogeneity before drawing any inferences.

TABLE 2

Pearson Correlation Coefficients (p-values) for Lease Model Variables during the Period 1995–2006

Variables	*Lease*	*AccQual*	*Nodiv*	*Unrated*	*Rating*	*Size*	*Loss*	*MTR*
Lease	1	−0.351	0.399	0.369	0.298	−0.532	0.471	−0.387
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
AccQual		1	−0.301	−0.233	−0.191	0.398	−0.398	0.318
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nodiv			1	0.363	0.482	−0.552	0.501	−0.417
				(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unrated				1	NA	−0.658	0.312	−0.192
						(0.001)	(0.001)	(0.001)
Rating					1	−0.568	0.495	−0.309
						(0.001)	(0.001)	(0.001)
Size						1	−0.644	0.492
							(0.001)	(0.001)
Loss							1	−0.695
								(0.001)
MTR								1

Correlations with *Rating* only apply to firms that are rated by the S&P.
See Table 1 for variable definitions.

The fourth and fifth columns in Table 3 partition the SDC subsample firms by whether the firm's bank loans contain the capital expenditure restriction. Firms that use this restrictive covenant tend to be smaller, lease less, face more financing constraints, and have lower performance. These firms also have smaller loans with longer maturities.

Multivariate Results

Panel A of Table 4 provides the results of the OLS estimation of Equation (1) for the comprehensive sample of manufacturing firms. Consistent with H1, we find that firms with worse accounting quality (i.e., low *AccQual*) have a greater propensity to lease assets. This result is both statistically and economically significant. For example, a standard deviation change in *AccQual* (standard deviation = 0.583) is associated with a 7.1 percent change in the mean *Lease* (mean = 0.436).[25]

As for control variables, results on *Nodiv*, *Unrated*, *Size*, *Loss*, and *MTR* are consistent with the reasoning that firms with more financial constraints have a greater propensity to lease assets. For comparative purposes, we repeat the analysis for the comprehensive sample on our lead lender subsample and report the results for this subsample in Panel B of Table 4. The results on both our test and control variables are quite similar across these two samples, suggesting that any differences between our comprehensive sample of Compustat manufacturing firms and the subsample of these firms with syndicated loans do not affect our lease-versus-buy findings.

[25] While not reported, we find that the VIF of our accounting quality measure is 1.25, so our finding that financial reporting quality influences leasing activities is not driven by multi-collinearity.

TABLE 3

Descriptive Statistics (means) for Sample Firms Partitioned on Whether the Firm Uses a Capital Expenditure Restriction in Bank Loans and the Lead Lenders' Ownership of Bank Loans for the Period 1995–2006

	High Lead Arrangers' Ownership	Low Ownership (t-stat for difference)	With Capital Expenditure Restrictions	Without Restrictions (t-stat for difference)
Lease	0.458	0.316 (10.58)***	0.371	0.325 (3.40)***
AccQual	0.071	0.302 (−9.94)***	0.223	0.249 (−1.02)
Nodiv	0.728	0.315 (14.29)***	0.556	0.391 (4.82)***
Unrated	0.888	0.359 (20.52)***	0.497	0.493 (0.13)
Rating[a]	3.802	3.097 (6.19)***	3.792	3.002 (12.17)***
Size	3.814	7.424 (−32.14)***	6.158	6.632 (−3.75)***
Loss	0.423	0.183 (14.00)***	0.319	0.226 (5.16)***
MTR	0.157	0.213 (−7.45)***	0.185	0.208 (−2.90)***
Lead Own	89.624	24.218 (69.44)***	NA	NA
Loan	16.364	19.433 (−41.35)***	18.091	18.326 (−2.13)**
Maturity	3.367	3.869 (−12.80)***	3.770	3.591 (5.16)***
n	499	498	374	491

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

[a] The mean values of this variable only apply to firms that are rated by the S&P.

Variable Definitions:
Lead Own = weighted average of the lead arrangers' ownership of the bank loan issued in the sample period;
Loan = natural of log of the amount of bank loan issued in the sample period; and
Maturity = natural of log of the maturity (in months) of bank loan issued in the sample period.

See Table 1 for definitions of other variables.

We present the results on our tests of the effect of banks' due diligence incentives on the lease-versus-buy decision in Table 5. In the first column of results, we report the determinants of the leasing propensity for firms where the lead lender holds a relatively larger proportion of the loan (i.e., *Lead Rank* = 1). In the second column of results, we report the determinants of the leasing propensity for firms where the lead lender holds a relatively smaller proportion of the loan (i.e., *Lead Rank* = 0).

Focusing on the determinants to lease for firms where the lead lender has a stronger incentive to monitor, we find that smaller firms, firms that suffer a loss, and firms that are not paying

TABLE 4

Coefficients (and t-statistics) from OLS Estimations of the Determinants of the Extent of Firms' Expenditures on Operating Leases

Panel A: Complete Sample

Variables	Predicted Sign	Coefficients	t-stats
Intercept	+/−	0.442	10.18***
AccQual	−	−0.053	−4.00***
Nodiv	+	0.048	3.35***
Unrated	+	0.057	1.67*
Rating	+	0.003	0.34
Size	−	−0.022	−7.26***
Loss	+	0.094	3.75***
MTR	−	−0.143	−2.88***
n		3,033	
R^2		0.3382	

Panel B: Syndicated Loan Sample

Variables	Predicted Sign	Coefficients	t-stats
Intercept	+/−	0.467	7.37***
AccQual	−	−0.074	−3.32***
Nodiv	+	0.036	2.21**
Unrated	+	0.028	0.73
Rating	+	0.003	0.32
Size	−	−0.021	−4.32***
Loss	+	0.085	2.66***
MTR	−	−0.102	−1.86*
n		1,314	
R^2		0.2414	

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.
All t-statistics are based on clustering at the three-digit industry level.
See Table 1 for variable definitions.

dividends are more likely to lease. These results suggest that firms facing financing constraints are more likely to lease, consistent with Eisfeldt and Rampini (2009) and Sharpe and Nguyen (1995). We also find that *AccQual* has a significant negative coefficient, indicating that worse accounting quality is associated with a greater leasing propensity for firms with a high lead arranger ownership in bank loans.

For firms where the lead lender has less monitoring incentives, we again find that *AccQual* has a significant negative coefficient. Consistent with the results reported in the first column, lower accounting quality is associated with a greater leasing propensity. We conduct a Chi-squared test on the coefficient on accounting quality across these two regressions to test H2. We find that the coefficient on accounting quality is statistically larger for firms that have lenders with fewer monitoring incentives. Thus, when information asymmetry problems are likely to be larger, accounting quality plays a relatively more important role in the leasing decision.

TABLE 5

Coefficients (and t-statistics) from Endogenous Switching Models Estimating How Monitoring by Banks, Measured by Lead Lenders' Ownership, Affects the Effect of Accounting Quality on Lease Expenditures

		Lead Rank = 1	*Lead Rank* = 0
Variables	**Predicted Sign**	**Coefficients (z-stats)**	**Coefficients (z-stats)**
Intercept	+/−	0.455 (1.75) *	0.402 (4.78)***
AccQual	−	−0.053 (−2.22) **	−0.211 (−6.48)***
Nodiv	+	0.073 (3.11) ***	0.043 (2.26)**
Unrated	+	0.044 (0.17)	0.063 (1.29)
Rating	+	0.006 (0.10)	0.011 (0.83)
Size	−	−0.020 (−2.73) ***	−0.011 (−1.45)
Loss	+	0.090 (1.98) **	0.053 (1.24)
MTR	−	−0.253 (−2.26) **	−0.015 (−0.21)
n		472	492

Differences in the impact of accounting quality on the lease-versus-buy decision

H0: $AccQual_{(Lead\ Rank=1)} = AccQual_{(Lead\ Rank=0)}$

$\chi^2\ (1) = 15.42$

p-value = 0.0001

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

See Table 1 for variable definitions.

We report the results of our tests of H3 on the effects of direct monitoring via the capital expenditure covenant on the leasing propensity in Table 6. The first column reports the results on the determinants of the leasing propensity for firms with investment-restriction covenants. The second results column reports the determinants of the leasing propensity for firms without investment-restriction covenants.

We find that for the subset of firms with investment covenants, the firms that are unrated, or are rated and have relatively poor ratings have a greater propensity to lease assets. Similarly, for the subset of firms without investment covenants, we find that smaller firms and firms that have suffered losses have a greater propensity to lease assets. Similar to the results discussed above, these findings support the findings in Eisfeldt and Rampini (2009) and Sharpe and Nguyen (1995) that firms that are more likely to have financing constraints are more likely to lease.

We find that accounting quality is an important determinant of leasing propensity for firms without investment covenants and is not a significant determinant of the leasing propensity for firms with these covenants. A Chi-squared test confirms that the effect of accounting quality on leasing propensity is significantly different across these regressions. These results support our

TABLE 6

Coefficients (and t-statistics) from Endogenous Switching Model Estimation of the Effect of Capital Expenditure Restrictions on Firms' Leasing Propensity

		Inv_Cov = 1	*Inv_Cov* = 0
Variables	**Predicted Sign**	**Coefficients (z-stats)**	**Coefficients (z-stats)**
Intercept	+/−	−0.602	0.364
		(−2.66)***	(5.18)***
AccQual	−	−0.005	−0.104
		(−0.15)	(−4.51)***
Nodiv	+	0.038	0.133
		(1.31)	(0.66)
Unrated	+	0.841	0.084
		(4.61)***	(1.24)
Rating	+	0.234	0.018
		(4.77)***	(0.93)
Size	−	−0.018	−0.013
		(−1.89)*	(−2.40)**
Loss	+	0.053	0.114
		(0.89)	(2.75)***
MTR	−	−0.184	−0.027
		(−1.38)	(−0.35)
n		373	483

Differences in the impact of accounting quality on the lease-versus-buy decision

H0: $AccQual_{(Inv\_Cov=1)} = AccQual_{(Inv\_Cov=0)}$

$\chi^2(1) = 5.66$

p-value = 0.017

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

Variable Definition:

Inv_Cov = indicator variable that equals 1 if the firm ever includes a capital expenditure restriction provision in its bank loans during the sample period.

See Table 1 for definition of other variables.

prediction in H3 that accounting quality is less important in leasing decisions when other mechanisms, like capital expenditure covenants, limit the ability of firms to engage in capital spending.

For completeness, we report the results of our "selection" regressions in Table 7. Although we do not develop hypotheses about the determinants of lead arranger's monitoring incentives or of investment covenants, we find that many of the explanatory variables included in the models are statistically significant. Of particular importance are the coefficients on the variables that serve as exogenous instruments. We find that large lead arranger ownership decreases with loan amounts and maturities, while capital expenditure provisions increases with loan amount and maturities. While not a direct test of "exogeneity," these results suggest that our exogenous instruments explain variation in investment covenants, and lead-ownership holdings.

TABLE 7

Coefficients (and z-statistics) from Probit Models Examining the Determinants of the Lead Lenders' Ownership and Use of Capital Expenditure Provisions for the Period 1995–2006

		Dependent Variable: *Lead Rank*	Dependent Variable: *Inv_Cov*
Variables	**Prediction**	**Coefficients (z-stats)**	**Coefficients (z-stats)**
Intercept	+/−	23.211	−6.531
		(10.98)***	(−5.82)***
AccQual	−	0.055	0.104
		(0.25)	(0.82)
Nodiv	+	−0.041	0.077
		(−0.27)	(0.71)
Unrated	+	−1.228	3.396
		(−1.73)*	(6.15)***
Rating	+	−0.250	0.984
		(−1.28)	(6.52)***
Size	−	−0.146	−0.104
		(−2.76)***	(−2.25)**
Loss	+	0.391	0.486
		(1.19)	(2.35)**
MTR	+/−	1.111	0.051
		(1.96)**	(0.11)
Loan	+/−	−1.132	0.010
		(−10.81)***	(1.78)*
Maturity	+/−	−0.314	0.440
		(−2.43)**	(4.28)***
n		964	856
Pseudo R^2		0.6732	0.1442

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

All z-statistics are based on clustering at the three-digit industry level.

Variable Definitions:

Lead Rank = indicator variable equals 1 if the average of the lead arrangers' ownership of the firm's bank loans exceeds the median of all sample firms, and 0 otherwise; and

Inv_Cov = indicator variable that equals 1 if the firm ever includes a capital expenditure restriction provision in its bank loans during the sample period.

See Table 1 for variable definitions.

VII. ROBUSTNESS CHECKS

Industry Composition

We focus on the manufacturing industry to avoid the heterogeneity in financing, leasing, and financial reporting choices across various industries. Untabulated results show that our findings continue to hold in analyses that include all industries and control for fixed industry effects.

Lease Measure

The main tabulated results are based on an approach that averages all variables across the sample period, where *Lease* is a stock measure. To show that our results are not driven by this

specific design, we also conduct a panel data analysis, where we measure accounting quality by averaging the absolute value of discretionary accruals over the three years prior to the measurement year of our leasing propensity variable.

Consistent with Sharpe and Nguyen (1995) and Eisfeldt and Rampini (2009), we also define *Lease* using a flow measure in the panel data framework. The flow measure of *Lease* is defined as the minimum lease payment (lagged Compustat #96) divided by the sum of lease and capital expenditures (Compustat #128). Table 8 shows that this alternative measure in our panel data analysis generates results consistent with our main approach, firms with higher accounting quality are less likely to lease than to buy.

We also replace the scalar in our model with the sum of total assets and capitalized lease obligations (Compustat # 6 + lagged Compustat #96 × 10). We continue to observe a significantly negative association between accounting quality and leasing. Finally, Tobit estimation including firms with missing lease data (about 8 percent of firms) produces similar results.

Accounting Quality Measure

In addition to using the principal-component approach to measuring accruals quality, we also derive measures based on the underlying variables, and include those measures in the model. Our

TABLE 8

Coefficients (and t-statistics) from OLS Estimations of the Determinants of the Extent of Firms' Expenditures on Operating Leases for the Period 1995–2006 Estimated Using an Annual Regression and Flow Measures of Lease Proclivity

Variables	Predicted Sign	Coefficients	t-stats
Intercept	+/−	0.522	19.68***
AccQual	−	−0.042	−5.98***
Nodiv	+	0.070	6.54***
Unrated	+	0.056	2.27**
Rating	+	0.005	0.69
Size	−	−0.029	−14.69***
Loss	+	0.030	4.14***
MTR	−	−0.135	−5.50***
n		15,055	
R^2		0.3021	

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

All t-statistics are based on clustering at the firm level. Results on time fixed effects are omitted.

Variable Definitions:

Lease = lease expenditure (lagged Compustat #96) divided by the sum of capital and lease expenditures (Compustat #128 + lagged Compustat #96);

AccQual= average of first principal component of three accrual-based accounting quality metrics based on Bharath et al. (2008) over three years prior to the year *Lease* is measured (described in the Appendix);

Nodiv = indicator variable equal to 1 if the firm does not pay a dividend (Compustat #21), and 0 otherwise;

Unrated= indicator variable equal to 1 if the firm is not rated by the S&P (Compustat #280), and 0 otherwise;

Rating = S&P ratings when the firm is rated by the S&P. Following Sharpe and Nguyen (1995), the rating is coded 1 if the S&P rating is between AAA and A−, 2 if the rating is between A+ and A−, 3 if the rating is between BBB+ and BBB−, and 4 if the rating is between BB+ and D;

Size = decile ranking of average of the natural log of sales (Compustat #12);

Loss = indicator variable for loss year, where the indicator variable equals 1 when the Compustat #123 is less than 0; and

MTR = marginal tax rates based on John Graham's simulation technique, as reported on his website.

results continue to hold when we use these underlying variables. To ensure that the decision to use cross-sectional estimation is not affecting our results, we also use a time-series measure of the extent of current accruals mapping into future, current, and future operating cash flows (Dechow and Dichev 2002) as a metric for accounting quality. The results of using this time-series measure are consistent with the results of using cross-sectional models.

Similarly, we change the scalar used in the accounting quality measures to be consistent with the scalar used on the independent variable, and to ensure that differences in the accounting treatment for leases are not driving our results. If we use total assets as the scalar, we find results similar to those that we tabulate.

Lead Rank Measure

To test the sensitivity of our results to our measure of lead ranks, we examined two alternative cutoffs. Setting this variable equal to 1 if the average of the lead arrangers' ownership of the firm's bank loans exceeds either the 33rd or the 66th percentile of all sample firms, and 0 otherwise, produces similar results.[26] The results of these alternative specifications suggest that our results are not driven by our median cutoff.[27]

Financial Constraints

Finally, when the firm has lower free cash flow, information problems are more likely to affect the firm's investment behavior through their effect on external financing. We interact accounting quality proxies with an indicator variable for high free cash flow. We find that the importance of accounting quality affecting leasing activities decreases with free cash flow, consistent with our prediction.

VIII. CONCLUSION

This study extends prior research on the relation between financial reporting quality and firms' investment and financing choices by investigating the effect of financial reporting quality on the lease-versus-buy decision. Our research design complements prior studies that focus exclusively on asset purchases when examining the effect of accounting quality on firm's investment decisions. We find that the leasing propensity is declining in accounting quality. This finding suggests that although previous research shows that firms with poor accounting quality buy fewer assets, they lease more of their assets, indicating that poor accounting quality may not result in lower total investment.

To distinguish between two alternative explanations of this finding, we also examine the mechanisms that lenders use to mitigate the agency problems arising from information asymmetries. Only if better accounting quality and leasing are two alternative mechanisms for addressing financial constraints would we expect the negative relation between accounting quality and leasing to be mitigated by other mechanism that reduce information asymmetry and agency costs. We find that accounting information is less important in the decision to lease when lenders' incentives to conduct due diligence are higher. The importance of financial reporting quality in the lease-versus-buy decision is also decreasing in the existence of capital expenditure covenants. Although the relation between leasing and accounting quality might be indicative of a propensity of firms with otherwise poor accounting quality to engage in off-balance-sheet financing, our results are more consistent with accounting quality and leasing serving as substitute mechanisms for reducing financing constraints.

---

[26] Similar to Sufi (2007), we find consistent results using a Herfindahl index of bank loan ownership.

[27] We also interacted a continuous lead-ownership measure with our accounting quality measure. The *AccQual* coefficient magnitude decreases with lead-ownership, consistent with the dichotomize approach.

Overall, our results highlight the importance of considering leasing when studying the effects of financial reporting quality on investments. Prior research suggests that firms with poor accounting quality have financial constraints, inhibiting their ability to invest in the periods in which they are liquidity constrained. Our results extend this line of research, suggesting that while low-quality accounting firms are likely to forgo purchasing assets, they are more likely to lease assets when they are liquidity constrained. Our analyses also contribute to the literature examining the importance of accounting information in firms' investment decisions by considering how banks' private information and monitoring can substitute for accounting information.

APPENDIX

Construction of *AccQual*

ABACC1: The absolute value of current discretionary accruals calculated based on Teoh et al. (1998). The model is estimated annually for each Fama and French (1997) industry group and each industry-year regression requires at least 20 observations. Based on Teoh et al. (1998), we first estimate the following regression to get the estimated coefficients (variables are defined below):

$$\frac{Current\_Accc}{Total\_Inv} = \gamma_1 \frac{1}{Total\_Inv} + \gamma_2 \frac{\Delta \text{Re}\, v}{Total\_Inv} + \eta.$$

The second step calculates the absolute value of discretionary accruals as:

$$\left| \frac{Current\_Acc}{Total\_Inv} - \hat{\gamma}_1 \frac{1}{Total\_Inv} - \hat{\gamma}_2 \frac{(\Delta \text{Re}\, v - \Delta AR)}{Total\_Inv} \right|.$$

ABACC2: The absolute value of total discretionary accruals calculated based on Dechow et al. (1995). The model is estimated annually for each Fama and French (1997) industry group and each industry-year regression requires at least 20 observations. We first estimate the following regression to get the estimated coefficients (variables are defined below):

$$\frac{Total\_Acc}{Total\_Inv} = \alpha_1 \frac{1}{Total\_Inv} + \alpha_2 \frac{\Delta \text{Re}\, v}{Total\_Inv} + \alpha_3 \frac{PPE}{Total\_Inv} + \varepsilon.$$

The second step calculates the absolute value of discretionary accruals as:

$$\left| \frac{Total\_Acc}{Total\_Inv} - \hat{\alpha}_1 \frac{1}{Total\_Inv} - \hat{\alpha}_2 \frac{(\Delta \text{Re}\, v - \Delta AR)}{Total\_Inv} - \hat{\alpha}_3 \frac{PPE}{Total\_Inv} \right|.$$

ABACC3: The absolute value of total current accruals calculated based on Dechow and Dichev (2002). The model is estimated annually for each Fama and French (1997) industry group and each industry-year regression requires at least 20 observations. ABACC3 is the absolute value of the estimated residual from the following model.

$$\frac{Current\_Acc}{Total\_Inv} = \theta_0 + \theta_1 \left(\frac{CFO}{Total\_Inv} \right)_{t-1} + \theta_2 \left(\frac{CFO}{Total\_Inv} \right)_{t} + \theta_3 \left(\frac{CFO}{Total\_Inv} \right)_{t+1} + v,$$

where:

Current_Acc= earnings before extraordinary items–cash flow from operating activities–depreciation (Compustat #123 − #308 + #14);[28]
Total_Acc = earnings before extraordinary items–cash flow from operating activities (Compustat #123 − #308);
Total_Inv = the sum of PP&E and capitalized lease expenditure (Compustat #8 + lagged Compustat #96 multiplied by 10);
CFO = cash flow from operating activities (Compustat #308);
ΔRev = change in sales (Compustat #12);
ΔAR = change in accounts receivables (Compustat #2).
PPE = property, plant, and equipment (Compustat #8).

We standardize each of the three measures by subtracting the mean and dividing by the standard deviation, and then extract the first principal component from the three proxies.[29] Our accounting quality variable (*AccQual*) is thus defined as the average of the following metric over the sample period 1995–2006:

$-1 * [\text{standardizedABACC1} * 0.38538 + \text{standardizedABACC2} * 0.38371 + \text{standardizedABACC3} * 0.32568]$.[30]

REFERENCES

Ball, R., R. Bushman, and F. Vasvari. 2008. The debt-contracting value of accounting information and loan syndicate structure. *Journal of Accounting Research* 46: 247–287.

Bharath, T., J. Sunder, and S. Sunder. 2008. Accounting quality and debt contracting. *The Accounting Review* 83: 1–28.

Biddle, G., and G. Hilary. 2006. Accounting quality and firm-level capital investment. *The Accounting Review* 81: 963–982.

———, ———, and R. Verdi. 2009. How does financial reporting quality improve investment efficiency? *Journal of Accounting and Economics* 48: 112–131.

Bushman, R., J. Piotroski, and A. Smith. 2005. Capital allocation and timely accounting recognition of economic losses: International evidence. Working paper, The University of Chicago.

———, A. Smith, and F. Zhang. 2008. Investment-cash flow sensitivities are really investment-investment sensitivities. Working paper, The University of North Carolina.

Christie, A. A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting and Economics* 9: 231–258.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77: 35–59.

———, W. Ge, and C. Schrand. 2009. Understanding earnings quality: A review of the proxies, their determinants and their consequences. Working paper, University of California, Berkeley.

———, R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–225.

Diamond, D. 1984. Financial intermediation and delegated monitoring. *The Review of Economic Studies* 51: 393–414.

[28] In robustness checks, we adjust the current accruals by excluding change in prepaid rental expenses to accommodate differences in accounting for operating leases versus asset purchase, and this adjustment does not affect the results.

[29] The first principal component explains that, on average, about 80 percent of the corresponding cross-sectional sample variance and only the first eigenvalue is greater than 1. We therefore conclude that the first principal component captures most of the common variation among the three proxies.

[30] We multiply by −1 to make the metric higher for firms with better accounting quality. The Pearson correlations between the first principal component and the three accounting quality proxies are very high, ranging from −80 percent to −95 percent.

Eisfeldt, L., and A. Rampini. 2009. Leasing, ability to repossess, and debt capacity. *Review of Financial Studies* 22: 1621–1657.

Fama, E. 1985. What's different about banks? *Journal of Monetary Economics* 15: 29–36.

———, and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43: 153–193.

Fazzari, S., R. G. Hubbard, and B. Petersen. 1988. Investment and finance reconsidered. *Brookings Papers on Economic Activity* 1: 141–195.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2004. Cost of equity and earnings attributes. *The Accounting Review* 79: 967–1010.

Holthausen, R., and R. Verrecchia. 1988. The effect of sequential information release on the variation of price changes in an intertemporal multi-asset market. *Journal of Accounting Research* 26: 82–106.

Hope, O. K., W. Thomas, and D. Vyas. 2009. Transparency, ownership, and financing constraints: An international study using private firms. Working paper, University of Toronto.

Kaplan, S. N., and L. Zingales. 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? *The Quarterly Journal of Economics* 112: 169–215.

———, and ———. 2000. Investment-cash flow sensitivities are not valid measures of financing constraints. *The Quarterly Journal of Economics* 115: 707–712.

Lee, L. 1978. Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. *International Economic Review* 19: 415–434.

Lara, J. M., B. G. Osma, and F. Penalva. 2009. Conditional conservatism and firm investment efficiency. Working paper, Universidad Carlos III de Madrid.

McNichols, M. 2002. Discussion of the quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77: 61–69.

Nini, G., D. Smith, and A. Sufi. 2009. Creditor control rights and firm investment policy. *Journal of Financial Economics* 92: 400–420.

Sharpe, S., and H. Nguyen. 1995. Capital market imperfections and the incentive to lease. *Journal of Financial Economics* 39: 271–294.

Smith, C., and M. Wakeman. 1985. Determinants of corporate leasing policy. *The Journal of Finance* 40: 895–908.

Sufi, A. 2007. Information asymmetry and financing arrangements: Evidence from syndicated loans. *The Journal of Finance* 62: 629–668.

Teoh, S. H., I. Welch, and T. J. Wong. 1998. Earnings management and the underperformance of seasoned equity offerings. *Journal of Financial Economics* 50: 63–99.

Verdi, R. 2006. Financial reporting quality and investment efficiency. Working paper, Massachusetts Institute of Technology.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1239–1271

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1239

Whistle-Blowing: Target Firm Characteristics and Economic Consequences

Robert M. Bowen
University of Washington
Andrew C. Call
The University of Georgia
Shiva Rajgopal
Emory University

ABSTRACT: We document the first systematic evidence on the characteristics and economic consequences of firms subject to employee allegations of corporate financial misdeeds. First, compared to a control group that avoided public whistle-blowing allegations, firms subject to whistle-blowing allegations were characterized by unique firm-specific factors that led employees to expose alleged financial misdeeds. Second, on average, whistle-blowing announcements were associated with a negative 2.8 percent market-adjusted five-day stock price reaction; this reaction was especially negative for allegations involving earnings management (−7.3 percent). Third, compared to a control group that exhibits similar characteristics, firms subject to whistle-blowing allegations were associated with further negative consequences including earnings restatements, shareholder lawsuits, and negative future operating and stock return performance. Finally, whistle-blowing targets exposed by the press were more likely to make subsequent improvements in corporate governance. Our results suggest whistle-blowing is far from a trivial nuisance for targeted firms, and on average, appears to be a useful mechanism for uncovering agency issues.

We acknowledge helpful comments on an earlier version of this paper from the editor, two anonymous referees, Alex Edwards, Jane Jollineau Kennedy, Allison Koester, Carol Marquardt, Devin Shanthikumar, and workshop participants at the 2008 Accounting and Ethics Conference at The University of Oklahoma, the 2007 FARS Midyear Meeting, the 2007 AAA Annual Meeting, INSEAD, London Business School, and The University of Washington. We thank Larry DuCharme for sharing his lawsuit data with us. We also acknowledge financial support from the Accounting Development Fund, the PricewaterhouseCoopers Professorship, and Julius A. Roller Professorship, all at the University of Washington Foster School of Business.

This paper is the recipient of the University of Oklahoma Price College of Business 2008 Glen McLaughlin Award for the best unpublished research paper in accounting ethics.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: July 2008
Accepted: December 2009
Published Online: June 2010

Keywords: *whistle-blowing; financial misdeeds; earnings management; corporate governance.*

Data Availability: *All data are available from public sources identified in the paper.*

I. INTRODUCTION

The purpose of our study is to identify characteristics of firms subject to whistle-blowing allegations, examine the economic consequences of such allegations, and thus shed light on the role of employee whistle-blowers in uncovering corporate financial misdeeds. Notable anecdotal evidence suggests that whistle-blowers can make a difference. For example, two whistle-blowers, Cynthia Cooper and Sherron Watkins, played significant roles in exposing accounting frauds at WorldCom and Enron, respectively, and were named as the 2002 "persons of the year" by *TIME* magazine. In response to Enron, WorldCom, and other scandals, Congress passed the Sarbanes-Oxley Act (SOX) in July 2002, which, in part, made it unlawful for companies to take punitive action against employees who disclose "questionable accounting or auditing matters" (SOX Section 806, codified as title 15 U.S.C., § 78f(m)(4)). Under the whistle-blower provisions of SOX, employees who disclose improper financial practices receive greater protection from discrimination (Title 18 U.S.C., § 1514A(a)(1)). SOX also required firms listed on a U.S. stock exchange to set up a hotline enabling whistle-blowers to communicate anonymously to the board of directors (*Economist* 2006). Further, the Securities and Exchange Commission is considering a whistle-blower program to encourage individuals to identify potential violations of federal securities laws (SEC 2009).

In contrast to these anecdotes suggesting that whistle-blowers help uncover financial improprieties, skeptics argue that (1) whistle-blowers often misjudge the situation and indulge in trivial or frivolous complaints (Miceli and Near 1992), (2) "Machiavellian" whistle-blowers who have an axe to grind lodge baseless allegations (Gobert and Punch 2000), and (3) ineffective workers misuse their protected "whistle-blower" status to avoid discharges or disciplinary proceedings (Schmidt 2003).

Despite the recent prominence of whistle-blowing events, we are not aware of prior work in the accounting and finance literatures that attempts to resolve the conflict between these perspectives. We examine the characteristics of firms subject to publicly available whistle-blowing allegations and the economic consequences of whistle-blowing events for target firms. We analyze 81 employee whistle-blowing allegations related to financial misconduct between 1989 and 2004 reported in the financial press (Press sample) and 137 instances of employee whistle-blowing between 2002 and 2004 obtained from records of the U.S. government's Occupational Safety and Health Administration office (OSHA), the agency responsible for handling complaints of discrimination from whistle-blowers who report SOX-related allegations (OSHA sample). We label these wide-ranging allegations as "external financial" whistle-blowing events. They are "external" because they are made known outside of the organization; we generally cannot observe cases of whistle-blowing that are resolved internally. They are "financial" because they include alleged earnings management, improper disclosure, insider trading, price-fixing in product markets, tax fraud, or violations of securities regulation. We investigate four related research questions:

Q1: What operating, financial, governance, industry, and regulatory characteristics differentiate firms subject to external financial whistle-blowing from other firms that avoid publicly disclosed whistle-blowing allegations?

Q2: What is the *immediate* stock market reaction to an external financial whistle-blowing announcement? Do these stock price reactions vary with the nature of the allegation?

Q3: What are the *subsequent* economic consequences for firms subject to external financial

whistle-blowing events? Do these firms experience more earnings restatements, more shareholder lawsuits, and/or poorer future operating and stock return performance?

Q4: Do firms respond to whistle-blowing allegations by improving their governance structure?

To investigate our first research question (Q1), we identify variables that potentially differentiate firms subject to public whistle-blowing actions. We argue that certain firm characteristics are associated with the probability that an employee will voice an allegation of financial impropriety to someone outside the organization. We compare whistle-blowing samples to two control groups: (1) a broad set of firms not subject to whistle-blowing allegations, and (2) firms "tarnished" by either a shareholder lawsuit or an earnings restatement (suggesting some level of financial misdeeds) that were able to avoid an external whistle-blowing allegation.

Compared to other tarnished firms, both Press and OSHA firms were more likely to be large, growing firms with strong recent stock return performance that had recently suffered employee layoffs. In addition, Press firms had relatively weak external monitoring and OSHA firms were relatively well regarded, had weak internal monitoring, and were more likely to be firms where the whistle-blowing employee had the opportunity to share in the proceeds of any fraud-related settlement with the U.S. government. Overall, we interpret our evidence as consistent with whistle-blowing targets being systematically different from other tarnished firms that avoided public whistle-blowing allegations of financial wrongdoing.

Our second research question (Q2) addresses the immediate stock market reaction to whistle-blowing allegations that are reported in the press. We find the average market-adjusted five-day stock return surrounding the announcement in the financial press was −2.8 percent. This reaction was especially negative if the allegation involved earnings management (−7.3 percent).

Our third research question (Q3) investigates subsequent economic consequences for firms targeted by whistle-blowing allegations, including restatements, lawsuits, and poor operating and stock return performance. Our research design uses a propensity-score matched control sample that includes firms with similar characteristics as the whistle-blowing targets to control for the fact that whistle-blowing is not an exogenous event and that negative consequences following whistle-blowing allegations could arise from many of the same underlying economic forces that contribute to the allegations themselves.

Compared to a propensity-score matched control sample, we find that whistle-blowing targets identified by the press were more likely to restate their earnings and be subject to shareholder litigation in the three years following the whistle-blowing allegation. In addition, we find evidence suggesting whistle-blowing targets (both Press and OSHA) suffered lower operating and stock return performance than did their control firm counterparts. Thus, whistle-blowers appeared to expose firms that (1) engaged in financial misconduct and (2) had previously unknown agency problems—both of which resulted in negative future consequences for the firm.[1] Evidence of such negative *ex post* consequences is consistent with the implicit premise in SOX that whistle-blowers can disclose valuable new information about firms' agency problems.

Our fourth research question (Q4) examines firms' responses to whistle-blowing allegations. We find that, compared to a matched control sample, firms subject to public whistle-blowing allegations (Press sample) were more likely to reduce the size of the board, reduce the proportion of insiders on the board, have less busy directors, and replace their CEO, whereas firms subject to less publicized whistle-blowing allegations (OSHA firms) did not respond to these allegations by making governance changes.

[1] Agency problems include corporate financial misdeeds that result from conflicts of interest between management and stakeholders such as shareholders and creditors.

Overall, our evidence suggests that whistle-blowing is an important mechanism for discovering information about agency problems at firms. On balance, our findings are inconsistent with claims by critics that whistle-blowing complaints are frivolous, misleading, or unreliable. Our research also provides indirect evidence on the efficacy of SOX Section 806 as a mechanism to uncover agency issues at firms. Interestingly, public whistle-blowing events are relatively infrequent, and a majority of the whistle-blowing allegations we identify are not directly related to the types of fraudulent behavior that precipitated the SOX legislation.

From a broader perspective, we contribute to an emerging stream of literature in corporate governance related to how information about financial wrongdoing reaches the stock market. In related work, Dyck et al. (2008) investigate who reveals information about corporate fraud. Using a sample of firms subject to shareholder class action litigation, they find that the most important sources of information are employees (19 percent of the cases), media (16 percent), and analysts (15 percent). Unlike Dyck et al. (2008), our focus is on (1) characteristics of firms targeted by employee whistle-blowers, (2) the economic consequences of whistle-blowing revelations, and (3) firms' responses via governance changes subsequent to such whistle-blowing events. In addition, Miller (2006) investigates the role of the press in revealing fraud, but does not indicate that employees are often named as the source of the articles. Furthermore, Miller (2006) identifies firms that are, *ex post*, known to have committed financial wrongdoing, and tracks these firms to determine the role of the press in uncovering these known cases of wrongdoing. In contrast, we identify firms *alleged* to have engaged in financial wrongdoing, and examine whether these firms ultimately suffered negative consequences.

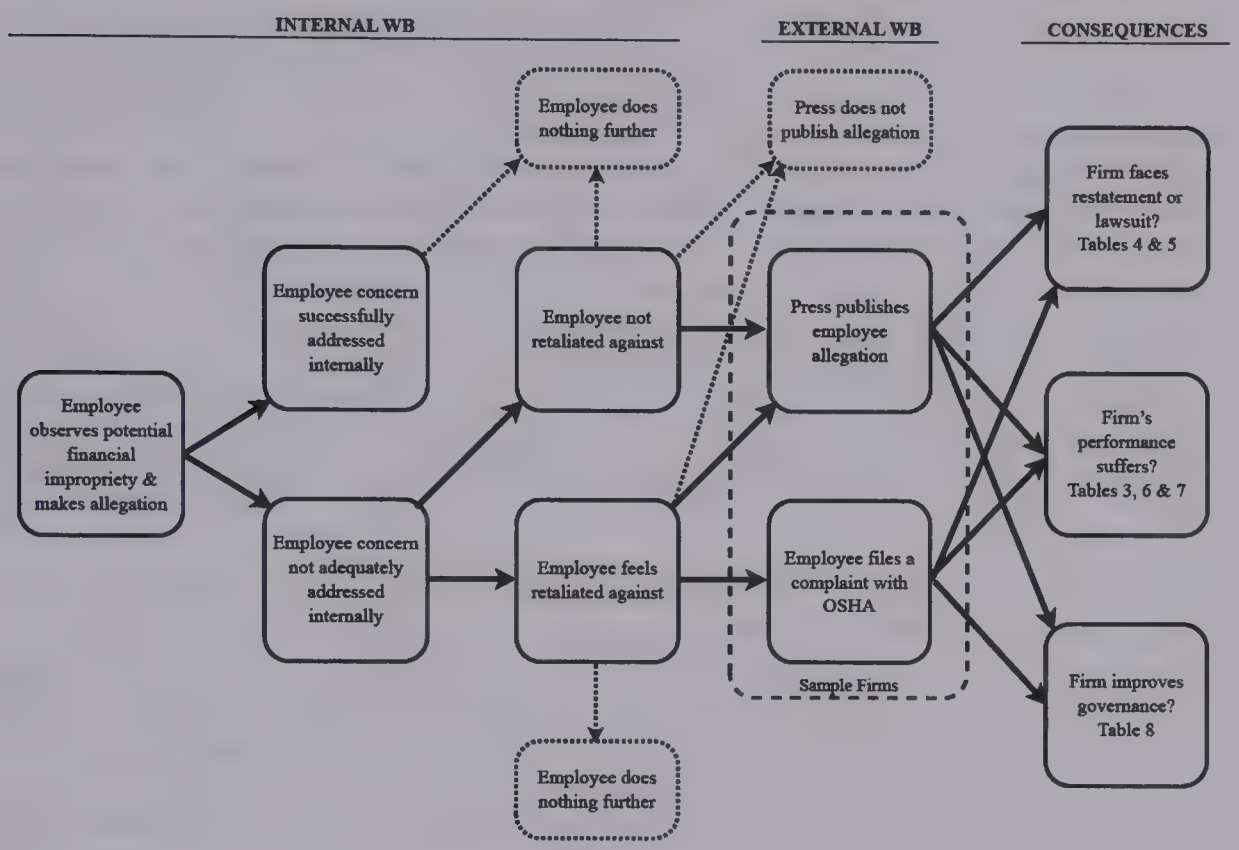
Section II discusses financial whistle-blowing and introduces predictions about firm-specific factors likely to be associated with whistle-blowing events. Section III describes the sample and Section IV presents evidence on the predictions discussed in Section II. Section V presents evidence on stock price reactions surrounding whistle-blowing allegations and also provides evidence on longer-term consequences of whistle-blowing actions. Section VI reports evidence on whether firms improve governance after whistle-blowing allegations. Section VII provides a summary of findings and contributions.

II. BACKGROUND AND CHARACTERISTICS OF FIRMS SUBJECT TO WHISTLE-BLOWING ALLEGATIONS

Financial Whistle-Blowing—Internal versus External

Miceli and Near (1985) define whistle-blowing as the disclosure by either former or current employees of alleged illegal, immoral, or illegitimate practices that are under employer control. Figure 1 summarizes outcomes of potential employee whistle-blowing allegations and related consequences for the targeted firm. Whistle-blowing can be either internal or external. Internal whistle-blowing involves informing relevant members of the firm (e.g., management, internal audit) about potential wrongdoing. External whistle-blowing, on the other hand, involves either an internal allegation becoming public or the whistle-blower directly disclosing an allegation outside the organization (e.g., to the media or a regulatory agency such as OSHA). We focus on external whistle-blowing because instances of internal whistle-blowing generally are not reported in the media and public data on such cases generally are not available. Moreover, we concentrate only on whistle-blowing related to financial practices.

FIGURE 1
Employee Whistle-Blowing Allegations: Sequence of Possible Actions and Consequences



Congressional Action to Protect Whistle-Blowers

Largely in response to the Enron and WorldCom scandals, Congress passed the Sarbanes-Oxley Act (SOX) in July 2002.[2] SOX, in part, sought to provide whistle-blowers greater legal protection. In addition to creating civil remedies for retaliation against employees in fraud cases, SOX made it a felony to take any action that is harmful to any person in retaliation for providing information about a federal crime to law enforcement officials. Thus, a retaliatory firing or demotion of a whistle-blowing employee would constitute a crime.

Whistle-Blowing is Frivolous, Misused, Misleading, and/or Unreliable

Some survey evidence (Brickey 2003; Zingales 2004) suggests that the personal cost to the whistle-blower can be high.[3] Given that whistle-blowers likely anticipate these costs, their allega-

[2] Enron and WorldCom represent high-profile examples of internal whistle-blowing that became public. Our general conclusions are not affected when we delete Enron and WorldCom from the sample, i.e., our results are not driven by these prominent examples.

[3] Zingales (2004) summarizes evidence on the consequences of whistle-blowing for individual employees. In a 1992 survey of 1,500 federal workers, 25 percent of employees reported that they experienced verbal harassment and intimidation; 20 percent were shunned by co-workers and managers; 18 percent were assigned to less desirable duties; and 11 percent were denied a promotion. A 1998 survey of 448 emergency physicians paints an even worse picture: 23 percent of those who complained about an issue reported having been fired or threatened with termination. Brickey (2003) reports that a random review of 200 whistle-blower complaints filed with the National Whistleblower Center in 2002 found that about half of the complainants said they were fired after they reported misconduct. The remaining complain-

tions should be credible. However, some critics suggest that whistle-blowers often misjudge the situation and lodge frivolous complaints to garner attention and publicity (Near and Miceli 1996). Indeed, in June 2005, the French government's Data Protection Authority refused to allow the installation of anonymous whistle-blower hotlines, saying that such lines "were disproportionate to the objectives sought with the risks of slanderous denunciations" (*Economist* 2006). Furthermore, troublemakers could potentially misuse their whistle-blower status to avoid being suspended or dismissed from employment. For example, Anechiarico and Jacobs (1996) argue that several municipal employees in New York City chose to blow the whistle in order to fall under state protection as a pre-emptive measure to being fired. Hence, given the enormous trouble that can result from disciplining a whistle-blower, supervisors and managers might find it easier to tolerate an unproductive and ineffective employee rather than dismiss or demote the employee. Gobert and Punch (2000) argue that an employee might even be excluded from necessary salary cuts to avoid the perception that such a cut is retaliation against the whistle-blower. On average, if whistle-blower complaints are frivolous, misused, misleading, and/or unreliable, then we should observe no significant economic consequences to a whistle-blowing action. Having said that, we acknowledge that even the taint of an allegation could potentially lead to some negative consequences for the firm.

Characteristics of Firms Subject to Whistle-Blowing Allegations: Opportunities and Incentives for Committing Financial Misdeeds

In this section, we argue that whistle-blowing allegations are related to management's opportunity and incentives to engage in financial wrongdoing and are a function of employees' incentives to voice their allegations outside the organization. At the outset, we acknowledge that some of these hypothesized factors are exploratory and descriptive. Given how little published research exists in the area of whistle-blowing, a solid theory is not available on why whistle-blowing occurs and what its consequences may be. Hence, we believe our descriptive approach is appropriate at this early stage of the literature's development. Factors we expect to explain external whistle-blowing follow.

Capital Market Pressure

In recent years, firms have faced heightened capital market pressure to deliver sustained earnings growth (Graham et al. 2005). Such pressures have likely created incentives for some managers to aggressively boost earnings, either via earnings management or via other questionable practices such as overbilling customers or improper disclosure of material financial events. We predict that firms subject to capital market pressure are more likely to experience a whistle-blowing event. To capture capital market pressure, we rely on three variables (described in detail in the Appendix) to proxy for (1) the need to access capital (*FIN_NEED*), (2) merger and acquisition activity (*M&A*), and (3) the amount of potential earnings management (*ABSDACC*). We compute the averaged scaled rank of these three proxies and label the averaged rank as *CM_PRESSURE*.

Growth

Firms that are rapidly growing are more likely to outgrow their controls. Responsibility for overall decision-making is typically spread across many individuals, and as a firm grows and responsibility is spread over more individuals, each individual has less information and authority

ants had been subjected to other retaliatory action such as on-the-job harassment or discipline. A survey by another watchdog group, the Government Accountability Project, found that about 90 percent of whistle-blowers are subjected to reprisals or threats.

to stop wrongdoing (Baucus and Near 1991). We predict that growth firms are more likely to be subject to a whistle-blowing event. We proxy for high-growth environments by using the scaled rank of sales growth (*GROWTH*) over the three years prior to the whistle-blowing event.

Past Performance

We predict that firms with strong past stock market performance are more likely to be a target of whistle-blowing for two reasons. First, these firms are more newsworthy to the media. Second, an employee might be more likely to expose financial misdeeds perceived to enable the firm to artificially achieve strong performance. We measure the prior performance of the firm (*PAST_PERF*) as the scaled rank of stock returns over the year before the whistle-blowing event.

Firm Reputation

We predict that a highly regarded firm is more likely to attract the ire of a disgruntled employee and the resulting allegation is more likely to be newsworthy. The external whistle-blower or the media outlet is likely to get more attention from exposing wrongdoing at a highly regarded firm.[4] We proxy for firm reputation based on whether the firm is listed either as a "Most Admired" firm or a "Best Place to Work For" by *Fortune* magazine (*REPUTATION*) in any of the five years prior to the whistle-blowing event.[5]

Unclear Internal Communication Channels

We hypothesize that the employee's motivation to resort to external whistle-blowing is higher when channels for raising concerns within the company are unclear. This situation likely occurs in at least two cases. First, Rothschild and Miethe (1999) suggest that bureaucratic and undemocratic work environments are likely to experience higher levels of *external* whistle-blowing because of significant barriers to effective *internal* whistle-blowing. Miceli and Near (1994) state that it is common for employees to first report problematic activities to management and, if the issue is ignored, then resort to reporting the problem to external audiences such as the media. We assume older firms (*AGE*) are likely to be more bureaucratic and, hence, more likely to be subject to a whistle-blowing action.

Second, King (1999) argues that geographical distance and multiple industrial segments in a firm make communication channels less clear. In particular, an employee who has to report a misdeed to a boss he or she hardly knows and who is geographically distant might feel uncomfortable raising sensitive issues (Miceli and Near 1994; Beresford et al. 2003). Such unclear internal communication channels increase the likelihood the employee will resort to external whistle-blowing. We measure industrial and geographical concentration as the Herfindahl-Hirschman index in the year prior to the whistle-blowing event based on revenues for each of the firm's industrial and geographic segments, respectively (*IND_CONC* and *GEO_CONC*). We expect the probability of external whistle-blowing to decrease with the industrial and geographic concentration of the firm's business segments. Our empirical measure for unclear communication channels (*COMMUNICATION*) is the averaged scaled rank of *AGE*, *IND_CONC*, and *GEO_CONC*.

Employee Downsizing

Luthans and Sommer (1999) find that following employee downsizing, employees experience a decline in both loyalty and commitment to the firm. We predict that employees, especially

[4] Of course, it is possible that a highly regarded firm is more likely to avoid organizational misconduct or have processes in place that resolve employee allegations before they are disclosed publicly.

[5] Note that the "Best Places" list is only available from 1998, while the "Most Admired" list is available throughout our entire sample period.

former employees who have been let go, are more likely to make public allegations following layoffs. Further, layoffs can increase the level of animosity between the firm and existing employees, and if existing employees perceive their job as being less secure, the potential cost of blowing the whistle decreases. We proxy for downsizing (*DOWNSIZE*) by using the scaled rank of the average percentage change in the number of employees over the three years prior to the whistle-blowing event.

Monetary Incentive to Report Financial Misdeeds (Qui Tam)

In certain instances, whistle-blowers have the potential to benefit financially by uncovering corporate misdeeds. Under the Federal False Claims Act, citizens who help uncover fraud in government contracts or government programs may be awarded a portion of the recovered funds. Since these whistle-blower provisions went into effect in 1986, the government has recovered over $15.6 billion in settlements and judgments from these *qui tam*-related cases. Of this amount, over 16 percent of the recovered funds have been distributed to the individuals who helped bring the fraud to light (TAF 2009).

The majority of *qui tam* cases fall into one of two industries: healthcare and defense contracting. Over 70 percent of the funds recovered as part of the *qui tam* provisions relate to matters in which the Department of Health and Human Services is the primary agency, while over 15 percent of the recovered funds pertain to matters in which the Department of Defense is the primary agency. We argue that the *qui tam* provisions included in the Federal False Claims Act provide employees in the healthcare and defense industries with increased incentives to blow the whistle. We proxy for these *qui tam* incentives and set *QUITAM* equal to 1 if the firm is in a healthcare industry (two-digit SIC code equal to 80) or if the firm has been on the Department of Defense's "100 Companies Receiving the Largest Dollar Volume of Prime Contract Awards" list in any of the three years prior to the whistle-blowing event.

Firm Size

The media is generally more interested in information about firms that are large (Miller 2006). It is also possible that regulators (e.g., OSHA) are similarly more likely to pay attention to information about large firms. As a result, employees of large firms are likely to be more confident that their allegations will have an audience, and we therefore predict that employees of large firms will be more likely to make allegations of financial wrongdoing. Our proxy for firm size (*SIZE*) is the scaled rank of sales in the year prior to the whistle-blowing event.

Internal Control Weaknesses

Internal controls weaknesses are defined as "a significant likelihood that a material misstatement of the annual or interim financial statements will not be prevented or detected" (PCAOB 2004). We argue that employees are more likely to resort to external whistle-blowing when the firm has poor internal controls. To capture the extent of a firm's internal control weaknesses (*ICW*), we use the scaled rank of the fitted value from a model of the determinants of internal control weaknesses, as outlined by Doyle et al. (2007) and described in the Appendix.

External Monitoring

Managers are more likely to engage in financial misconduct when the quality of external monitoring is poor. To capture such external monitoring we use two empirical proxies: (1) the presence of block-holders (*BLOCK*) and (2) public pension funds ownership (*PP*). Consistent with prior literature (e.g., Cremers et al. 2005), we argue that firms for which *BLOCK* and *PP* are low are characterized by relatively poor external monitoring. We use the averaged scaled rank of *BLOCK* and *PP* measured in the year prior to the whistle-blowing event and label the averaged

rank *EXT_MONITOR*. We expect a negative association between whistle-blowing allegations and monitoring by powerful shareholders.

Governance

Managers are more likely to indulge in wrongdoing in the presence of weak corporate governance and strong private incentives. To capture the quality of internal governance, we obtain board data from the IRRC database, and use the averaged scaled rank of seven empirical proxies (described in detail in the Appendix) related to (1) the CEO's private incentives and (2) overall board structure. We label the averaged rank as *GOVERNANCE*.

Model for Assessing Characteristics of Firms Subject to Whistle-Blowing Allegations

Next, we estimate a logistic model and regress the presence or absence of a whistle-blowing allegation on the above-mentioned proxies. We introduce industry fixed effects (based on two-digit SIC codes other than for SIC code 80 covered by *QUITAM*) and year fixed effects as controls for potential unspecified factors. Model 1 (where firm and year subscripts are suppressed) links whistle-blowing events to managerial incentives to commit financial misdeeds and the whistle-blowers' likelihood of coming forward:

$$\begin{aligned} Pr(\textit{Whistle-blowing}) = {} & \beta_0 + \beta_1 CM\_PRESSURE + \beta_2 GROWTH + \beta_3 PAST\_PERF \\ & + \beta_4 REPUTATION + \beta_5 COMMUNICATION + \beta_6 DOWNSIZE \\ & + \beta_7 QUITAM + \beta_8 SIZE + \beta_9 ICW + \beta_{10} EXT\_MONITOR \\ & + \beta_{11} GOVERNANCE + \text{Industry} + \text{Year} + \text{error} \end{aligned} \quad (1)$$

where *Pr*(*Whistle-blowing*) is the probability of a whistle-blowing event for this binary dependent variable, i.e., 1 = a whistle-blowing event and 0 = none. Detailed definitions for the independent variables are in the Appendix.

III. WHISTLE-BLOWING SAMPLES AND CONTROL FIRMS

Whistle-Blowing Samples

Our whistle-blower sample comes from two sources, as illustrated in Figure 1. The first source, labeled the Press sample, consists of 81 whistle-blowing events drawn from a Lexis-Nexis search of every combination of the following two groups of search terms (1) "whistle," "whistle-blowing," "whistleblower," "whistle-blower," and (2) "financial," "accounting," "reporting," "fraud," "accounting fraud" over the calendar years 1989 to 2004. Note that the Press sample label does not imply that employee whistle-blowers directly informed the press. These are merely instances of whistle-blowing reported in the media.

The second source, labeled the OSHA sample, consists of 137 whistle-blowing discrimination cases gathered from written requests to the Occupational Health and Safety Administration (OSHA).[6] While OSHA handles other varieties of whistle-blower allegations besides those who

[6] While the vast majority of OSHA cases were not resolved in favor of the employee whistle-blower (Moberly 2007; *New York Times* 2009), this does not negate the importance or credibility of the allegations. First, OSHA's rulings were decisions about whether the whistle-blower was *discriminated* against—*not* rulings on whether the alleged financial impropriety was taking place. It is entirely possible that the underlying allegation of impropriety was credible and important, even though OSHA determined the employee was not treated unfairly. Second, Moberly (2007) finds that one of the reasons employee whistle-blowers have not been successful at winning claims is because "OSHA tended to misapply Sarbanes-Oxley's burden of proof regarding causation, to the substantial detriment of employees." Thus, despite the limited success employee whistle-blowers have had in their discrimination cases filed with OSHA, it remains unclear whether these employees were raising concerns about legitimate cases of financial malfeasance, e.g., a *New York Times* (2009) editorial argues that "Most often, the reason for the dismissal has nothing to do with the merits of the

claim to have uncovered fraud and file for protection under SOX, we asked for and received data related only to financial whistle-blowing under the purview of SOX.[7] These cases represent whistle-blowers' complaints of discrimination in the workplace (e.g., demotion, termination) for having voiced allegations of financial impropriety. In essence, the OSHA sample reflects the very whistle-blowers SOX Section 806 was intending to protect. While we are unable to observe whistle-blowing allegations that did not result in alleged employee discrimination (because these allegations were not reported to OSHA) and did not otherwise become public, each OSHA case represents a previously internal allegation of financial wrongdoing being made public. Figure 1 highlights the external whistle-blowing allegations we are able to observe.

The Press sample differs from the OSHA sample in several ways. First, the Press sample likely reflects the media's bias in covering more visible, important, or sensational whistle-blowing allegations. In contrast, the OSHA sample is likely to be less biased but is more likely to include relatively frivolous complaints. Second, Callahan and Dworkin (1994) find that misdeeds reported to the media are likely to involve larger sums of money than those reported to other outlets. Third, observations in the Press sample, by definition, represent a clear public revelation of the whistle-blowing allegation whereas the OSHA whistle-blowing data are less likely to be widely disseminated in local and national media given the information has to be obtained via a special request from a government agency. Fourth, we can ascertain the date on which the whistle-blowing allegation was made public based on the news story in the press; such event-dating is difficult with the OSHA sample. Hence, we conduct a study of the stock market reaction surrounding the whistle-blowing event only for the Press sample. However, to the extent possible, we examine the long-run consequences of both types of whistle-blowing events.

We start with an initial Press (OSHA) sample of 81 (137) unique firms.[8] Panel A of Table 1 presents an annual breakdown of the firms subject to whistle-blowing actions. The number of whistle-blowing actions in the Press sample is relatively stable at a median value of 4 per year, except for a peak of 23 events in 1998 and 13 events in 2002. Because the OSHA sample begins in August 2002, this year has only 12 events, followed by 62 events in 2003 and 63 events in 2004. Panel B of Table 1 lists the nature of whistle-blowing allegations in both samples. These diverse allegations range from clearly "earnings management" to potential earnings management (e.g., "overbilling" with the motivation to boost revenue) to other accounting decisions that affect disclosure (e.g., "improper disclosure" and "accounting irregularities") to other types of financial impropriety not directly related to financial reporting (e.g., "insider trading," "price-fixing" and "tax fraud").

Control Samples

We rely on two control samples. The purpose of the first control sample (labeled "broad control sample") is to benchmark our whistle-blowing firms against a broad set of firms that have not been publicly accused of financial improprieties by employee whistle-blowers. This sample consists of all firms that have non-missing CRSP and Compustat data for all of the independent variables listed in model (1). All observations that have non-missing data in the year before the whistle-blowing event for one of our sample firms are eligible for inclusion in the broad control

---

case." Our study provides direct evidence on this question.

[7] The Sarbanes-Oxley Act did not assign the job of protecting whistle-blowers to the Securities and Exchange Commission. Instead, the task went to the Labor Department's OSHA because the law's drafters reasoned that OSHA had a record of helping workers penalized for voicing concerns. For example, OSHA has the responsibility of enforcing 13 additional whistle-blower statutes covering truckers, airline mechanics, nuclear-power-plant employees, and others.

[8] The number of firms used in each empirical test is reported in the tables that follow and varies depending on data availability.

TABLE 1

Whistle-Blowing Allegations by Year and Type

Panel A: Whistle-Blowing Allegations by Year

	Press	OSHA		Press	OSHA
1989	1	NA	1997	2	NA
1990	0	NA	1998	23	NA
1991	1	NA	1999	2	NA
1992	3	NA	2000	5	NA
1993	4	NA	2001	4	NA
1994	5	NA	2002	13	12
1995	5	NA	2003	8	62
1996	5	NA	2004	0	63
				81	137

Panel B: Whistle-Blowing by Type of Allegation

	Press	%	OSHA	%
Earnings Management	24	29.6%	16	11.7%
Improper Disclosure	1	1.2%	4	2.9%
Overbilling	40	49.4%	3	2.2%
Insider Trading	3	3.7%	1	0.7%
Price Fixing	6	7.4%	0	0.0%
Tax Fraud	2	2.5%	0	0.0%
Securities Law Violations	0	0.0%	10	7.3%
Other Accounting-Related Allegations	1	1.2%	26	19.0%
Corporate Governance Issues	0	0.0%	3	2.2%
Fraud	0	0.0%	14	10.2%
Other	4	4.9%	18	13.1%
No Info	0	0.0%	42	30.7%
	81	100.0%	137	100.0%

sample at the first pass. We next impose the following filters: (1) no observation relating to any of the sample firms can be included in the broad control sample, and (2) no control firm can appear in the broad control sample more than once (i.e., no more than one firm-year observation for each control firm). Whenever we found multiple observations from the same firm in the control sample, we randomly selected just one firm-year.

One drawback to benchmarking the whistle-blowing sample to a broad sample of firms is that firms subject to a whistle-blowing event are inherently different from the average firm. Thus, we construct a second control group by comparing whistle-blowing firms to other firms tarnished by allegations of financial misdeeds that have *not* been exposed to a public whistle-blowing event. In particular, we identify firms that were subject to a lawsuit during the years 1989–2006 ("lawsuit sample") and firms that filed earnings restatements between 1997–2005 based on two GAO reports ("restatement sample"). We emphasize comparisons between our whistle-blowing samples and this second control group because they allow us to more directly assess differences between firms that face whistle-blowing allegations and other tarnished firms that avoid public allegations.

We rely on an updated version of the lawsuit sample of 1,119 usable firms compiled by DuCharme et al. (2004) covering the period April 1988 through May 2006 to identify firms that

were sued in connection with alleged financial misdeeds. The restatement sample is drawn from two lists of firms complied by the General Accounting Office (GAO 2002, 2006) through a Lexis-Nexis search with variations of the word "restate." Data limitations on several of the independent variables constrain the restatement sample to 1,304 firm-year observations.[9]

IV. CHARACTERISTICS OF FIRMS FACING EXTERNAL WHISTLE-BLOWING ALLEGATIONS (Q1)

In this section, we present evidence on the characteristics of the whistle-blowing targets versus the two control samples discussed in Section III. Given the Press and OSHA samples differ, we separately compare each whistle-blowing sample to the control samples.

Whistle-Blowing Firms versus the Broad Control Sample

In Panel A of Table 2, we compare whistle-blowing firms to the broad control sample of firms that have avoided external whistle-blowing allegations. We estimate multivariate model (1) separately for the Press sample in columns 1 and 2 and the OSHA sample in columns 3 and 4. Columns 1 and 3 present results without the external monitoring and internal governance variables because data restrictions on these variables severely restrict our sample sizes. Results in columns 2 and 4 include external monitoring and internal governance variables.

We find that, relative to the broad control sample, whistle-blowing firms in both Press and OSHA samples (1) were growing more rapidly (with statistically significant differences reported in columns 1, 2, and 3), (2) had relatively unclear internal communication channels (all four columns), (3) came from *QUITAM* environments where the whistle-blower had an opportunity to share in the proceeds of any fraud-related settlement with the U.S. Government (columns 1, 3, and 4), and (4) were larger (all four columns). Results in columns 2 and 4 suggest that whistle-blowing targets also had relatively weak external monitoring. We also find some evidence that capital market pressures were associated with whistle-blowing (columns 2 and 3). Two variables were only significant for the OSHA sample. OSHA firms were more likely to (1) have a strong reputation (columns 3 and 4), and (2) have undergone employee downsizing (column 3).

Whistle-Blowing Firms versus Lawsuit and Restatement Firms

In Panel B of Table 2, we compare the whistle-blowing sample to a combined sample composed of the lawsuit and restatement control samples.[10] Although both whistle-blowing firms and these control firms (lawsuit and restatement firms) are likely tarnished, only the whistle-blowing targets experienced a public allegation. Hence, our analysis in Panel B focuses on differences between firms where employees made external whistle-blowing allegations and lawsuit and restatement firms *where a public whistle-blowing allegation presumably could have been made.*

Compared to lawsuit and restatement firms, whistle-blowing firms (1) were growing more rapidly (significant in all four columns of Panel B), (2) were relatively more successful in their prior stock return performance (all four columns), (3) had experienced employee downsizing (columns 2, 3, and 4), and (4) were larger (all four columns).[11] In addition, OSHA firms (1) were more highly regarded (columns 3 and 4), (2) came from *QUITAM* environments (columns 3 and

[9] We eliminate firms that overlap the whistle-blowing treatment sample and the tarnished-firm control sample (lawsuits and restatements) to avoid any mechanical correlation between these samples.

[10] In an effort to capture the period of alleged financial malfeasance for our control firms, for lawsuit firms we calculate the average of all our continuous variables over the years of alleged wrongdoing. For restatement firms we calculate the average of all our continuous variables over the three years prior to the restatement announcement ($t-1$ through $t-3$). For comparability, we also use prior three-year averages for the continuous variables in our whistle-blowing sample.

[11] We acknowledge that the finding that whistle-blowing firms are characterized by relatively more successful prior stock return performance may be driven by the fact that lawsuits are often triggered by poor stock performance.

TABLE 2

What Firm Characteristics Lead Employees to Make Whistle-Blowing Allegations Public? (Q1)

Panel A: Whistle-Blowing Sample versus Broad Control Sample

$$\Pr(\text{whistle-blowing}) = \beta_0 + \beta_1 CM\_PRESSURE + \beta_2 GROWTH + \beta_3 PAST\_PERF + \beta_4 REPUTATION + \beta_5 COMMUNICATION + \beta_6 DOWNSIZE + \beta_7 QUITAM + \beta_8 SIZE + \beta_9 ICW + \beta_{10} EXT\_MONITOR + \beta_{11} GOVERNANCE + \text{Industry} + \text{Year} + \text{error} \quad (1)$$

		(1)		(2)		(3)		(4)	
		Press versus Control		Press versus Control		OSHA versus Control		OSHA versus Control	
Determinant	Pred. Sign	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
CM_PRESSURE	+	1.46	0.141	4.26	0.031**	2.41	0.007***	−0.64	0.663
GROWTH	+	2.40	0.013**	5.03	0.001***	1.25	0.022**	1.27	0.117
PAST_PERF	+	−0.84	0.869	−2.86	0.987	−0.56	0.854	−0.78	0.819
REPUTATION	+	0.29	0.344	0.47	0.328	1.46	0.001***	1.11	0.014**
COMMUNICATION	−	−2.66	0.006***	−2.93	0.054*	−1.23	0.061*	−2.03	0.053*
DOWNSIZE	−	−0.00	0.499	−0.52	0.359	−1.29	0.015**	−0.12	0.450
QUITAM	+	0.92	0.043**	0.39	0.319	1.33	0.012**	1.61	0.019**
SIZE	+	8.26	<0.001***	28.27	<0.001***	3.01	<0.001***	7.14	<0.001***
ICW	+	−0.86	0.874	−0.11	0.534	−0.27	0.699	0.97	0.106
EXT_MONITOR	−			−1.86	0.076*			−1.96	0.019**
GOVERNANCE	−			−1.01	0.234			1.05	0.771
Pseudo-R^2		50.6%		65.8%		39.0%		43.1%	
Likelihood Ratio Statistic		354.42	<0.001***	280.37	<0.001***	361.78	<0.001***	194.85	<0.001***
ROC Curve Statistic		0.944		0.983		0.942		0.943	
n		$n_{PRESS} = 67$ $n_{CONTROL} = 4{,}552$		$n_{PRESS} = 50$ $n_{CONTROL} = 1{,}282$		$n_{OSHA} = 95$ $n_{CONTROL} = 4{,}552$		$n_{OSHA} = 54$ $n_{CONTROL} = 1{,}282$	

(continued on next page)

Panel B: Whistle-Blowing Sample versus Tarnished (Combined Lawsuit and Restatement) Samples

$$\Pr(\text{whistle-blowing}) = \beta_0 + \beta_1 CM\_PRESSURE + \beta_2 GROWTH + \beta_3 PAST\_PERF + \beta_4 REPUTATION + \beta_5 COMMUNICATION + \beta_6 DOWNSIZE + \beta_7 QUITAM + \beta_8 SIZE + \beta_9 ICW + \beta_{10} EXT\_MONITOR + \beta_{11} GOVERNANCE + \text{Industry} + \text{Year} + \text{error} \quad (1)$$

		(1)		(2)		(3)		(4)	
		Press versus Lawsuit/Restate		Press versus Lawsuit/Restate		OSHA versus Lawsuit/Restate		OSHA versus Lawsuit/Restate	
Determinant	**Pred. Sign**	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**
CM_PRESSURE	+	0.03	0.499	−1.42	0.667	1.54	0.120	−3.63	0.937
GROWTH	+	1.93	0.087*	7.63	0.003***	2.02	0.003***	2.42	0.046**
PAST_PERF	+	3.08	0.009***	4.17	0.078*	3.98	<0.001***	5.19	0.001***
REPUTATION	+	1.16	0.200	0.77	0.354	1.13	0.023**	1.55	0.012**
COMMUNICATION	−	−1.97	0.163	−0.35	0.419	1.15	0.868	0.49	0.599
DOWNSIZE	−	−0.86	0.239	−4.92	0.032**	−2.18	<0.001***	−1.88	0.072*
QUITAM	+	0.12	0.457	−3.37	0.949	2.23	0.010***	3.96	0.006***
SIZE	+	5.62	0.001***	20.18	0.003***	3.19	<0.001**	4.90	0.022**
ICW	+	0.54	0.343	1.49	0.261	−0.37	0.700	−0.22	0.571
EXT_MONITOR	−			−2.37	0.163			−2.33	0.055*
GOVERNANCE	−			−7.53	0.034**			1.73	0.803
Pseudo-R^2		67.7%		79.5%		52.6%		59.6%	
Likelihood Ratio Statistic		265.19	<0.001***	211.44	<0.001***	371.65	<0.001***	186.74	<0.001***
ROC Curve Statistic		0.976		0.995		0.963		0.970	
n		$n_{PRESS} = 38$ $n_{LAWSUIT} = 2{,}397$		$n_{PRESS} = 30$ $n_{LAWSUIT} = 915$		$n_{OSHA} = 80$ $n_{LAWSUIT} = 2{,}397$		$n_{OSHA} = 43$ $n_{LAWSUIT} = 915$	

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

p-values are one-sided for variables with directional predictions. We report (1 − p) values for coefficients that assume a sign opposite to the one predicted.

All variables are as defined in the Appendix. Firm and year subscripts are suppressed.

4), and (3) had relatively weak external monitoring (column 4). Press firms appeared to have worse governance (column 2).

In general, a comparison of whistle-blowing firms to lawsuit and restatement firms suggests whistle-blowing firms were systematically different from other tarnished firms. These unique environmental factors encouraged whistle-blowers to make their allegations public and help explain why some firms were subject to whistle-blowing allegations while other firms that presumably could have been targeted by whistle-blowers did not experience such allegations.

V. IMMEDIATE (Q2) AND SUBSEQUENT (Q3) CONSEQUENCES OF WHISTLE-BLOWING ALLEGATIONS

Immediate Stock Market Reactions to Whistle-Blowing Announcements

Recall that we can compute stock market reactions only for the Press sample, as the dates on which the OSHA whistle-blowing events become public are not generally available. Panel A of Table 3 presents evidence on the stock market reaction to whistle-blowing events measured around five days, from the day before the announcement to three days after (day −1 through day +3). *RET* represents raw returns while *CAR* signifies market-adjusted returns. The mean five-day abnormal market-adjusted stock market reaction was −2.84 percent. The median abnormal reactions, while smaller in magnitude, were also significantly negative. Given that the median market capitalization of a Press sample firm was $7.3 billion (untabulated), market reactions following whistle-blowing allegations resulted in significant losses in shareholder value, on average. The frequency of negative reactions is also high, at 67.9 percent of the sample.[12]

These significant negative reactions suggest that, on average, announcements of whistle-blowing allegations were economically important for the targeted firm and were not immediately seen as frivolous by stock market participants. We consider subsequent economic consequences of these whistle-blowing allegations later in Section V.

Cross-Sectional Determinants of the Stock Market Reactions to Whistle-Blowing Announcements

In Panel B of Table 3, we attempt to explain the cross-sectional distribution of stock market reactions to a whistle-blowing event by including five factors that potentially affect investors' reaction to whistle-blowing allegations. First, we expect larger firms (*SIZE*) to suffer larger negative stock price reactions. Larger firms have deeper pockets and are more likely to be the subject of a subsequent class action lawsuit. Second, highly regarded firms have more credibility to lose from an external whistle-blowing action reported in the financial press. We measure a firm's reputation based on whether the firm is ranked as most admired or best to work for per *Fortune* magazine (*REPUTATION*). Third, awareness of whistle-blowing actions likely heightened after the passage of SOX, and we expect the stock market to react more negatively to a whistle-blowing action in the post-SOX era. We define a variable, *SOX*, set equal to 1 if the whistle-blowing action was reported in the press after the calendar year 2002. Fourth, we argue that allegations of earnings management and overbilling represent especially large agency issues, and we predict these allegations are likely to result in particularly negative stock market reactions. *EM* is set to 1 if the allegation pertains to earnings management (and 0 otherwise) and *OVERBILLING* is set to 1 if the allegation involves overbilling (and 0 otherwise).

Panel B of Table 3 presents evidence on the cross-sectional variation in stock price reactions around whistle-blowing allegations. Two of our predictions are consistently significant. We find that the stock market viewed whistle-blowing allegations relatively negatively for firms with

[12] We find qualitatively similar results when we examine three-day (day −1 through day +1) returns instead of five-day returns.

TABLE 3

What is the Immediate Stock Market Reaction to Whistle-Blowing Announcements? (Q2)

Panel A: Five-Day Market Reaction (n = 78)

	% negative	Mean	Q1	Median	Q3	p-value[a]	p-value[b]
RET	60.3%	−0.0246	−0.0478	−0.0125	0.0208	0.055*	0.033**
CAR	67.9%	−0.0284	−0.0592	−0.0147	0.0284	0.024**	0.013**

Panel B: Cross-Sectional Determinants of the Market Reaction

Model : $CAR = \beta_0 + \beta_1 SIZE + \beta_2 REPUTATION + \beta_3 SOX + \beta_4 EM + \beta_5 OVERBILLING + \varepsilon$

Independent Variables	Pred. Sign	Coefficient	t-statistic
Intercept	?	0.070	1.73*
SIZE	−	−0.000	−0.69
REPUTATION	−	−0.024	−0.31
SOX	−	−0.041	−0.87
EM	−	−0.143	−3.04***
OVERBILLING	−	−0.074	−1.62*
Adj. R^2		8.9%	
F-statistic		2.24	0.063*
n		64	

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

p-values are one-sided for variables with directional predictions. We report (1 − p) values for coefficients that assume a sign opposite to the one predicted.

[a] p-values are one-sided associated with t-statistics.

[b] p-values are one-sided associated with Wilcoxon Z-statistics.

Variable Definitions (where firm and year subscripts are suppressed):

RET = five-day cumulative raw returns beginning the day before the whistle-blowing allegation;
CAR = five-day cumulative abnormal returns beginning the day before the whistle-blowing allegation;
SIZE = total sales revenue for the year before the whistle-blowing event;
REPUTATION = an indicator variable coded as 1 if the firm has appeared on *Fortune*'s "Best Companies to Work For" list or on *Fortune*'s "Most Admired Companies" list in any of the five years prior to the whistle-blowing event, and coded as 0 otherwise;
SOX = an indicator variable equal to 1 if the allegation was made later after 2002, 0 otherwise;
EM = an indicator variable equal to 1 if the allegation is related to earnings management, 0 otherwise; and
OVERBILLNG = an indicator variable equal to 1 if the allegations related to overbilling, 0 otherwise.

alleged earnings management. Specifically, the coefficient on *EM* is −0.143 (t-statistic = −3.04) indicating that allegations related to earnings management resulted in an average incremental (overall) −14.3 percent (−7.3 percent) return over the five-day window. Allegations of overbilling indicate an average incremental (overall) −7.4 percent (−0.4 percent) five-day return. Hence, it appears that, on average, whistle-blowing events were far from trivial for targeted firms, particularly for firms accused of managing earnings.

Subsequent Consequences of Whistle-Blowing Events (Q3)

In this section, we explore four potential subsequent consequences of whistle-blowing events: (a) earnings restatements, (b) shareholder lawsuits, (c) future operating performance, and (d) future stock market performance. If a whistle-blowing event reveals financial reporting or agency

problems in the firm, then one would expect firms subject to such allegations to be more likely to file future earnings restatements, to be the target of shareholder lawsuits, and to suffer weak future operating and stock return performance.

We recognize the issue of endogeneity in evaluating the consequences of whistle-blowing events. In other words, when we assess future economic consequences such as operating or stock return performance, it is possible that whistle-blowing allegations arise from many of the same economic forces that contribute to subsequent problems after a whistle-blowing allegation. Hence, it would be inappropriate to treat the whistle-blowing event as exogenous, especially given that we attempt to model the factors associated with whistle-blowing in Section IV.

We control for the potential endogeneity between whistle-blowing allegations and future economic consequences by comparing whistle-blowing targets to a sample of control firms matched on the propensity to experience a whistle-blowing event (Armstrong et al. 2009; Francis and Lennox 2008). The primary benefit of using a control sample matched on propensity scores is that it allows us to compare whistle-blowing firms to a set of firms that are the same on all observable dimensions, thus allowing us to more clearly attribute any observed consequence to the whistle-blowing event itself, rather than to the firm characteristics associated with whistle-blowing allegations. From a practical standpoint, this matched sample also allows us to compare whistle-blowing firms to control firms along a variety of dimensions (e.g., restatements, lawsuits, operating performance, stock return performance) without the need for additional control variables. In essence, our propensity-score matching procedure directly controls for any relevant difference between whistle-blowing and control firms.

To identify the propensity-score matched control sample, we estimate model (1) for the whistle-blowing firms and the broad control sample (as outlined in Panel A of Table 2), and calculate a propensity score for each firm, which is the conditional probability of receiving the treatment effect (i.e., the probability that the firm experienced a whistle-blowing event) given all the observable data.[13] For each whistle-blowing target, we select two control firms with the closest propensity scores, and these firms constitute the propensity-score matched control sample. The result is that each whistle-blowing firm is matched to two firms that were not subject to whistle-blowing allegations but are similar along all other relevant dimensions.[14] Since the matched control sample is similar along every observable dimension, any difference in outcome between whistle-blowing and control samples can be attributed to the whistle-blowing event.

We measure "covariate balance" to ensure that the whistle-blowing and matched control samples are similar across all dimensions except the variable of interest (i.e., being the target of whistle-blowing allegations).[15] In general, untabulated results suggest that in comparing whistle-blowing firms to these control firms, we effectively control for the potential endogeneity linking whistle-blowing allegations and future consequences. However, we do find a significant difference in the *distribution* (but not the *mean*) of *SIZE* when comparing whistle-blowing and control samples. Therefore, to be conservative, we control for *SIZE* in subsequent tests.

---

[13] In order to maximize the number of observations available for the propensity score matching procedure, we estimate model (1) excluding *EXT_MONITOR* and *GOVERNANCE*.

[14] We choose two matched firms for each whistle-blowing target in an effort to maximize the number of firms in this control sample while ensuring that there is no statistical difference in the propensity scores between treatment and control firms.

[15] We formally assess covariate balance by testing whether the means and medians of the covariates used in model (1) differ between the whistle-blowing and controls samples. The untabulated results suggest that the propensity-score matched samples resemble the whistle-blowing samples along virtually all dimensions. Specifically, there is no significant difference in means of any of the covariates between whistle-blowing and matched control firms (for either the Press or OSHA samples). Furthermore, the Kolmogorov-Smirnov test fails to reject the null hypothesis of similar distributions for each of these variables, with the exception of *SIZE*.

Do Whistle-Blowing Announcements Lead to Subsequent Earnings Restatements? (Q3a)

To assess subsequent consequences (in this and the following lawsuit-based tests), we rely on the restatement and lawsuit samples discussed in Section III with one major exception. Unlike Section III, we do not eliminate whistle-blowing firms that experienced a restatement or lawsuit, as the purpose of this analysis is to document the existence of such a link should it exist.[16,17]

The results in Panel A of Table 4 indicate that whistle-blowing firms in the Press sample experienced significantly more subsequent restatements ($p = 0.013$) in the three years following a whistle-blowing action (17.9 percent) than propensity-score matched control firms (7.5 percent). Panel B of Table 4 reports that whistle-blowing firms in the OSHA sample also experienced more restatements (13.7 percent) than propensity-score matched control firms (9.5 percent), but the difference is not significant at conventional levels ($p = 0.141$). However, when we look only at the subset of future restatements that represent accounting irregularities (as defined by Hennes et al. [2008]), we find that OSHA firms were more likely to experience an accounting irregularity than their matched counterparts ($p = 0.053$, untabulated).

Because the whistle-blowing and propensity-score matched control samples had different *SIZE* distributions, we also controlled for *SIZE*. These results, reported in Panel C of Table 4, support our earlier conclusions. In sum, firms exposed to whistle-blowing events reported in the press were more likely to experience a subsequent restatement than were matched control firms with similar underlying characteristics.

Do Whistle-Blowing Announcements Lead to Subsequent Shareholder Lawsuits? (Q3b)

Table 5 provides evidence on whether whistle-blowing allegations were a precursor to shareholder lawsuits. The sample analyzed for this test again includes 67 Press firms and their 134 propensity-score matched control firms, and 95 OSHA firms along with their 190 matched control firms. The results in Panel A of Table 5 indicate that whistle-blowing firms in the Press sample experienced significantly more subsequent lawsuits ($p = 0.004$) following a whistle-blowing action (26.9 percent) than propensity-score matched control firms (11.9 percent). Panel B of Table 5 reports that whistle-blowing firms in the OSHA sample also experienced more lawsuits (11.6 percent) than propensity-score matched control firms (8.4 percent), but the difference is not significant at conventional levels ($p = 0.196$). Controlling for *SIZE* in Panel C yields similar conclusions. In summary, the evidence in Tables 4 and 5 suggests that a whistle-blowing event reported in the press can be an early warning indicator of future earnings restatements or shareholder lawsuits, and that whistle-blowing allegations are not trivial and/or baseless, on average.

Are Whistle-Blowing Events Associated with Relatively Poor Future Operating Performance? (Q3c)

In this and the following section, we evaluate whether whistle-blowing allegations serve as an early warning of negative future operating or stock price performance. The link between a whistle-blowing event and negative future performance can be motivated in at least three ways. First, given our earlier evidence that whistle-blowing can be a significant indicator of subsequent earn-

[16] We verified that the whistle-blowing announcement preceded the earnings restatement before the announcement entered the empirical analyses. However, note that the GAO report discloses the date of the restatement announcement, and not the years being restated. Data on the year being restated is generally patchy and unreliable. Hence, we did not pursue an investigation into the year for which the financial statements were restated.

[17] The time-windows and data selection for the restatement tests need some elaboration. We assume that whistle-blowing events increase the probability of a restatement announcement within the next three years (including the year of the allegation). Because the GAO restatement data are available only from 1997 to 2005, we are forced to delete observations that occur in and before 1993. Thus, the sample analyzed for these tests includes 67 Press firms and their 134 propensity-score matched control firms, and 95 OSHA firms along with their 190 propensity-score matched control firms.

TABLE 4

Long-Term Consequences: Are Whistle-Blowing Allegations Associated with Subsequent Restatements? (Q3a)

Panel A: Restatement Frequency—Press versus Propensity-Score Matched Sample

Press Sample			Propensity-Score Matched Sample			Significance[a]
# of Restatements	# of Firms	% with Restatements	# of Restatements	# of Firms	% with Restatements	p-value
12	67	17.9%	10	134	7.5%	0.013**

Panel B: Restatement Frequency—OSHA versus Propensity-Score Matched Sample

OSHA Sample			Propensity-Score Matched Sample			Significance[a]
# of Restatements	# of Firms	% with Restatements	# of Restatements	# of Firms	% with Restatements	p-value
13	95	13.7%	18	190	9.5%	0.141

Panel C: Logistic Regressions

$$RESTATE_{t+3} = \beta_0 + \beta_1 WB_t + \beta_2 SIZE_{t-1} + \varepsilon_{t+3}$$

Independent Variables	Pred. Sign	Press versus Propensity-Score Matched Sample		OSHA versus Propensity-Score Matched Sample	
		Coefficient	p-value	Coefficient	p-value
WB	+	1.010	0.016**	0.406	0.157
SIZE	?	0.000	0.470	0.000	0.621
Pseudo-R^2		4.1%		0.8%	
Likelihood Ratio Statistic		5.432	0.066*	1.467	0.480
ROC Curve Statistic		0.666		0.560	
n		$n_{PRESS} = 66$ $n_{CONTROL} = 124$		$n_{OSHA} = 94$ $n_{CONTROL} = 182$	

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.
p-values are one-sided for variables with directional predictions. We report (1 − p) values for coefficients that assume a sign opposite to the one predicted.

[a] p-values are associated with the Pearson Chi-square test statistic, and test for differences in the proportion of firms that experience a restatement.

Variable Definitions (where firm subscripts are suppressed):
$RESTATE_{t+3}$ = 1 if the firm experienced a restatement in the event year or in the three subsequent years, 0 otherwise;
WB_t = 1 if the firm experienced a whistle-blowing event, 0 otherwise; and
$SIZE_{t-1}$ = total sales revenue (Compustat annual #12) for the year before the whistle-blowing event.

ings restatements and lawsuits, it is perhaps natural to expect negative future performance subsequent to a whistle-blowing event. Second, managers may engage in defensive actions to mitigate negative publicity and other consequences of a whistle-blowing event. Such defensive efforts can distract managers from normal profit-seeking activities and siphon funds away from strategic investments. Third, a whistle-blowing event can hurt the firm's image with its stakeholders, result in a loss of confidence in management, increase the probability of management turnover, and/or increase the perception that additional problems are lurking.

TABLE 5

Long-term Consequences: Are Whistle-Blowing Allegations Associated with Subsequent Lawsuits? (Q3b)

Panel A: Lawsuit Frequency—Press versus Propensity-Score Matched Sample

Press Sample			Propensity-Score Matched Sample			Significance[a]
# of Lawsuits	# of Firms	% with Lawsuits	# of Lawsuits	# of Firms	% with Lawsuits	p-value
18	67	26.9%	16	134	11.9%	0.004***

Panel B: Restatement Frequency—OSHA versus Propensity-Score Matched Sample

OSHA Sample			Propensity-Score Matched Sample			Significance[a]
# of Lawsuits	# of Firms	% with Lawsuits	# of Lawsuits	# of Firms	% with Lawsuits	p-value
11	95	11.6%	16	190	8.4%	0.196

Panel C: Logistic Regressions

$$LAWSUIT_{t+3} = \beta_0 + \beta_1 WB_t + \beta_2 SIZE_{t-1} + \varepsilon_{t+3}$$

Independent Variables	Pred. Sign	Press versus Propensity-Score Matched Sample		OSHA versus Propensity-Score Matched Sample	
		Coefficient	p-value	Coefficient	p-value
WB	+	0.849	0.017**	0.234	0.293
SIZE	?	0.000	0.025**	0.000	0.308
Pseudo-R^2		6.2%		0.8%	
Likelihood Ratio Statistic		10.598	0.005***	1.484	0.476
ROC Curve Statistic		0.728		0.584	
n		$n_{PRESS} = 66$		$n_{OSHA} = 94$	
		$n_{CONTROL} = 124$		$n_{CONTROL} = 182$	

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

p-values are one-sided for variables with directional predictions. We report (1 − p) values for coefficients that assume a sign opposite to the one predicted.

[a] p-values are associated with the Pearson Chi-square test statistic, and test for differences in the proportion of firms that experience a lawsuit.

Variable Definitions (where firm subscripts are suppressed):
$LAWSUIT_{t+3}$ = 1 if the firm experienced a lawsuit in the event year or in the three subsequent years, 0 otherwise;
WB_t = 1 if the firm experienced a whistle-blowing event, 0 otherwise; and
$SIZE_{t-1}$ = total sales revenue (Compustat annual #12) for the year before the whistle-blowing event.

In analyzing the link between whistle-blowing events and future operating performance, we calculate cumulative return on assets (*ROA*) for three time windows following the whistle-blowing event ranging from one to three years (i.e., *ROA* for year +1; years +1 and +2; and years +1, +2, and +3). The data in Panels A and B of Table 6 indicate that both Press and OSHA firms report lower *ROA* in the year immediately following the whistle-blowing allegation. Beyond the first year, whistle-blowing firms did not report statistically lower *ROA*. In Panels C and D, we control for *SIZE*, finding similar results, except that Press firms also reported lower *ROA* in the two years after the whistle-blowing event.

TABLE 6

Long-Term Consequences: Are Whistle-Blowing Allegations Associated with Poor Future Operating Performance? (Q3c)

Panel A: Future ROA—Press versus Propensity-Score Matched Sample

	Means			Medians		
	Press	**Control**	**p-value[a]**	**Press**	**Control**	**p-value[b]**
ROA_{t+1}	0.020	0.049	0.015**	0.035	0.043	0.020**
	(58)	(101)		(58)	(101)	
ROA_{t+2}	0.061	0.095	0.106	0.081	0.094	0.201
	(53)	(81)		(53)	(81)	
ROA_{t+3}	0.099	0.128	0.255	0.109	0.098	0.501
	(44)	(59)		(44)	(59)	

Panel B: Future ROA—OSHA versus Propensity-Score Matched Sample

	Means			Medians		
	OSHA	**Control**	**p-value[a]**	**OSHA**	**Control**	**p-value[b]**
ROA_{t+1}	0.017	0.044	0.070*	0.035	0.048	0.030**
	(84)	(155)		(84)	(155)	
ROA_{t+2}	0.057	0.063	0.445	0.071	0.087	0.264
	(40)	(59)		(40)	(59)	
ROA_{t+3}	0.160	0.115	0.619	0.233	0.229	0.526
	(6)	(8)		(6)	(8)	

Panel C: OLS Regressions—Press versus Propensity-Score Matched Sample

$$ROA_{t+n} = \beta_0 + \beta_1 WB_t + \beta_2 SIZE_{t-1} + \varepsilon_{t+n}$$

Independent Variables	**Pred. Sign**	ROA_{t+1} **Coeff.**	ROA_{t+1} **p-value**	ROA_{t+2} **Coeff.**	ROA_{t+2} **p-value**	ROA_{t+3} **Coeff.**	ROA_{t+3} **p-value**
WB	–	−0.025	0.032**	−0.037	0.097*	−0.024	0.303
SIZE	?	0.000	0.040**	0.000	0.012**	0.000	0.140
Adj. R^2		3.1%		4.1%		0.3%	
F-statistic		3.41	0.036**	3.66	0.029**	1.13	0.327
n		$n_{PRESS} = 55$		$n_{PRESS} = 49$		$n_{PRESS} = 41$	
		$n_{CONTROL} = 96$		$n_{CONTROL} = 76$		$n_{CONTROL} = 56$	

Panel D: OLS Regressions—OSHA versus Propensity-Score Matched Sample

$$ROA_{t+n} = \beta_0 + \beta_1 WB_t + \beta_2 SIZE_{t-1} + \varepsilon_{t+n}$$

Independent Variables	**Pred. Sign**	ROA_{t+1} **Coeff.**	ROA_{t+1} **p-value**	ROA_{t+2} **Coeff.**	ROA_{t+2} **p-value**	ROA_{t+3} **Coeff.**	ROA_{t+3} **p-value**
WB	–	−0.035	0.029**	−0.021	0.317	0.018	0.457
SIZE	?	0.000	0.168	0.000	0.274	0.000	0.685
Adj. R^2		1.1%		−0.8%		−15.4%	
F-statistic		2.29	0.103	0.64	0.527	0.13	0.878
n		$n_{OSHA} = 82$		$n_{OSHA} = 39$		$n_{OSHA} = 6$	
		$n_{CONTROL} = 153$		$n_{CONTROL} = 58$		$n_{CONTROL} = 8$	

(continued on next page)

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

Coefficients (t-statistics) are reported. p-values are one-sided for variables with directional predictions. We report (1 − p) values for coefficients that assume a sign opposite to the one predicted.

[a] p-values are associated with t-statistics.

[b] p-values are associated with Wilcoxon z-statistics.

Variable Definitions (where firm subscripts are suppressed):

ROA_{t+n} = buy-and-hold return on assets, starting with *ROA* in the year after the event date, and ending *n* years after the whistle-blowing event;

WB_t = 1 if the firm experienced a whistle-blowing event, 0 otherwise; and

$SIZE_{t-1}$ = total sales revenue (Compustat annual #12) for the year before the whistle-blowing event.

Are Whistle-Blowing Events Associated with Relatively Poor Future Stock Price Performance? (Q3d)

Table 7 presents data on future stock return performance of firms subject to whistle-blowing allegations. In particular, we consider monthly returns over three years following the month in which the whistle-blowing allegation becomes public. Thus, this event window does not overlap with our immediate consequences tests reported earlier in Section V. Panel A of Table 7 reveals Press firms reported lower abnormal returns over the one, two, and three years following the whistle-blowing event. Panel B reports some evidence consistent with OSHA firms also reporting lower returns over the subsequent two or three years.

In Panels C and D of Table 7, we regress raw monthly returns RET on (1) *SIZE* because whistle-blowing firms and the propensity-score matched control firms did not exhibit the same distribution along this one dimension, and (2) the other three Fama-French factors, including momentum. We include a *WB* indicator equal to 1 for firms in the whistle-blowing sample, and 0 otherwise, which allows for a different intercept for whistle-blowing firms and control firms. The regressions are conducted using the firm-month as the unit of analysis. The variable of interest in these analyses is *WB*. If whistle-blowing allegations were associated with poor future stock return performance, then we ought to observe a negative coefficient on *WB*, after controlling for other factors influencing stock returns. The results reported in Panels C and D of Table 7 are generally consistent with the univariate results. Press firms exhibited lower returns in year +1 and year +2, while OSHA firms exhibited lower returns in year +2 and year +3. In general, we find that firms targeted by whistle-blowers suffered lower returns in subsequent years, which in turn suggests that whistle-blowing allegations were not frivolous and without merit, on average.

VI. DO WHISTLE-BLOWING TARGETS RESPOND BY CHANGING CORPORATE GOVERNANCE? (Q4)

In this section, we investigate our fourth research question (Q4): What actions (if any) do whistle-blowing targets take to improve corporate governance following the revelation of a whistle-blowing allegation? It is not clear whether firms will take meaningful actions to improve corporate governance following a whistle-blowing event. For example, a targeted firm can easily initiate cosmetic changes such as replacing one set of inside directors with another set of insiders. In extreme cases, a target firm might even fire the whistle-blower.

TABLE 7

Long-Term Consequences: Are Whistle-Blowing Allegations Associated with Poor Future Stock Return Performance? (Q3d)

Panel A: Future Returns—Press versus Propensity-Score Matched Sample

	Means			Medians		
	Press	Control	p-value[a]	Press	Control	p-value[b]
$ABNRET_{t+1}$	−0.118	−0.001	0.047**	−0.107	0.000	0.030**
	(60)	(101)		(60)	(101)	
$ABNRET_{t+2}$	−0.312	0.037	0.001***	−0.434	−0.106	0.001***
	(54)	(89)		(54)	(89)	
$ABNRET_{t+3}$	−0.261	0.177	0.019**	−0.292	−0.074	0.010***
	(52)	(84)		(52)	(84)	

Panel B: Future Returns—OSHA versus Propensity-Score Matched Sample

	Means			Medians		
	OSHA	Control	p-value[a]	OSHA	Control	p-value[b]
$ABNRET_{t+1}$	−0.068	−0.014	0.197	−0.112	−0.034	0.152
	(90)	(167)		(90)	(167)	
$ABNRET_{t+2}$	−0.123	−0.013	0.137	−0.249	−0.077	0.081*
	(82)	(150)		(82)	(150)	
$ABNRET_{t+3}$	−0.156	0.064	0.080*	−0.221	−0.078	0.276
	(77)	(136)		(77)	(136)	

Panel C: OLS Regressions—Press versus Propensity-Score Matched Sample

$$RET_{t+n} = \beta_0 + \beta_1 WB_t + \beta_2 MKT_{t+n} + \beta_3 SMB_{t+n} + \beta_4 HML_{t+n} + \beta_5 MOM_{t+n} + \varepsilon_{t+n}$$

Independent Variables	Pred. Sign	RET_{t+1}		RET_{t+2}		RET_{t+3}	
		Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Intercept	?	0.725	0.076*	0.653	0.011**	0.699	0.001***
WB	–	−0.950	0.070*	−0.530	0.091*	−0.370	0.130

(continued on next page)

Panel C: OLS Regressions—Press versus Propensity-Score Matched Sample

$$RET_{t+n} = \beta_0 + \beta_1 WB_t + \beta_2 MKT_{t+n} + \beta_3 SMB_{t+n} + \beta_4 HML_{t+n} + \beta_5 MOM_{t+n} + \varepsilon_{t+n}$$

Independent Variables	Pred. Sign	RET_{t+1} Coeff.	RET_{t+1} p-value	RET_{t+2} Coeff.	RET_{t+2} p-value	RET_{t+3} Coeff.	RET_{t+3} p-value
MKT	?	1.098	<0.001***	1.069	<0.001***	1.057	<0.001***
SMB	?	0.416	<0.001***	0.277	<0.001***	0.168	<0.001***
HML	?	0.501	<0.001***	0.568	<0.001***	0.550	<0.001***
MOM	?	−0.231	<0.001***	−0.241	<0.001***	−0.252	<0.001***
Adj. R^2		13.0%		14.1%		13.7%	
F-statistic		57.30	<0.001***	112.60	<0.001***	150.34	<0.001***
n		$n_{PRESS} = 691$ $n_{CONTROL} = 1,189$		$n_{PRESS} = 1,286$ $n_{CONTROL} = 2,104$		$n_{PRESS} = 1,851$ $n_{CONTROL} = 2,863$	

Panel D: OLS Regressions—OSHA versus Propensity-Score Matched Sample

$$RET_{t+n} = \beta_0 + \beta_1 WB_t + \beta_2 MKT_{t+n} + \beta_3 SMB_{t+n} + \beta_4 HML_{t+n} + \beta_5 MOM_{t+n} + \varepsilon_{t+n}$$

Independent Variables	Pred. Sign	RET_{t+1} Coeff.	RET_{t+1} p-value	RET_{t+2} Coeff.	RET_{t+2} p-value	RET_{t+3} Coeff.	RET_{t+3} p-value
Intercept	?	0.281	0.327	0.362	0.034**	0.390	0.018**
WB	–	−0.314	0.242	−0.519	0.046**	−0.478	0.030**
MKT	?	1.225	<0.001***	1.151	<0.001***	1.175	<0.001***
SMB	?	0.269	0.002***	0.346	<0.001***	0.403	<0.001***
HML	?	0.470	<0.001***	0.387	<.001***	0.318	<0.001***
MOM	?	−0.074	0.201	−0.074	0.080*	−0.156	<0.001***
Adj. R^2		10.9%		11.1%		12.9%	
F-statistic		74.73	<0.001***	137.67	<0.001***	212.73	<0.001***
n		$n_{OSHA} = 1,050$ $n_{CONTROL} = 1,967$		$n_{OSHA} = 1,939$ $n_{CONTROL} = 3,564$		$n_{OSHA} = 2,614$ $n_{CONTROL} = 4,546$	

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

(continued on next page)

Coefficients (t-statistics) are reported. p-values are one-sided for variables with directional predictions. We report (1 − p) values for coefficients that assume a sign opposite to the one predicted.

[a] p-values are associated with t-statistics.

[b] p-values are associated with Wilcoxon Z-statistics.

Variables Definitions (where firm subscripts are suppressed):

$ABNRET_{t+n}$ = buy-and-hold abnormal returns (equally weighted) for the subsequent n years, beginning with the monthly return immediately following the event date;
RET_{t+n} = buy-and-hold raw returns for the subsequent n years, beginning with the monthly return immediately following the event date;
WB_t = 1 if the firm experienced a whistle-blowing event, 0 otherwise;
MKT_{t+n} = the Fama-French market factor for the subsequent n years, beginning with the month immediately following the event date;
SMB_{t+n} = the Fama-French size factor for the subsequent n years, beginning with the month immediately following the event date;
HML_{t+n} = the Fama-French book-to-market factor for the subsequent n years, beginning with the month immediately following the event date; and
MOM_{t+n} = the Fama-French momentum factor for the subsequent n years, beginning with the month immediately following the event date.

Changes in Governance Variables

We hand-collect governance data for the Press and OSHA whistle-blowing samples from their 10-Ks and proxy statements.[18] To limit data collection to manageable levels, we focus on six aspects of governance, details of which have been discussed earlier in Section II: (1) the number of directors on the board (*DIRECTORS*); (2) the proportion of insiders on the board (*%INS_DIR*); (3) the proportion of the board composed of busy directors where busy is defined as three additional directorships (*%BUSY*); (4) the percentage attendance at board meetings (*%ATTEND*); (5) the proportion of firms where the CEO is also the chairman of the board (*CEO=COB%*), and (6) the proportion of firms with a new CEO (*NEW_CEO%*). We calculate three-year changes in these six governance variables from the year before the whistle-blowing event (year −1) to two years after the event occurs (year +2). Of the 81 (137) possible Press (OSHA) firms, we are unable to find 10-Ks and/or proxy statements for 24 (47) firms.

Since governance improvements were encouraged in the period following the accounting scandals and SOX, we also collect governance data for a sample of matched firms that were not subject to a whistle-blowing allegation. We select matched firms with the same four-digit SIC code, within the same stock exchange, and with net sales as close as possible to the whistle-blowing firm's net sales for the year before the whistle-blowing event.[19,20]

Table 8 reports the multivariate analyses that compare each governance change between the Press sample and their matched-firms counterparts. Control variables include: size (*SALES*), past stock market performance (*RET*), and the change in firm size over the corresponding three-year period ($\Delta SIZE$) to control for the possibility that changes in governance (e.g., smaller board size) were simply driven by changes in firm size. We find that, compared to the matched control sample, firms exposed to whistle-blowing allegations via the press were more likely to reduce the size of the board, reduce insider participation on the board, have less busy board members, and were more likely to replace the CEO.

In untabulated findings, we find no evidence that OSHA firms made changes in governance. Thus, on average, whistle-blowing targets exposed via the press appeared to change governance on four dimensions that exceeded the level of market-wide governance improvements occurring in the post-SOX era, while firms targeted by whistle-blowers via OSHA (a less public venue) did not. This suggests the visibility of the allegation played a role in firms' responses to whistle-blowing allegations.

VII. CONCLUSIONS

Whistle-blowing has received considerable attention in recent years after (1) whistle-blowers were responsible, in part, for revealing the accounting scandals at Enron and WorldCom, and (2) provisions in SOX were enacted to protect employee whistle-blowers. Still, little is known about the nature of firms that are subject to whistle-blowing and whether such allegations are economically significant events with meaningful consequences for the targeted firms. Unlike other research

---

[18] Note that we hand-collect proxy data to assess subsequent governance changes, rather than use the IRRC data used in Section II, in order to maximize the number of useable observations in this analysis.

[19] If we cannot find a match firm within the same four-digit SIC code, then we relax the matching criteria to include firms with the same three-digit SIC code, and if necessary, the same two-digit SIC code.

[20] We employ a matched control sample in this analysis to ensure that any observed governance changes are not the result of market-wide improvements in governance (rather than improvements specific to whistle-blowing targets). Our matching criteria are consistent with this concern. We do not match on propensity scores (as with our consequences analyses) because propensity matches are not necessarily aligned in calendar time, and because we want to examine whether whistle-blowing targets improved governance relative to a broad cross-section of firms. We acknowledge that comparing whistle-blowing firms to the average firm results in a somewhat lower benchmark than would comparing them to a propensity-score matched sample. However, we feel this is appropriate because our sample period already represents a period of improved governance, such that the average firm was likely making governance improvements.

TABLE 8

Do Whistle-Blowing Allegations Lead to Subsequent Changes in Corporate Governance? (Q4)

Multivariate Analysis—Press Sample:

$$\Delta GOV_{t+2} = \beta_0 + \beta_1 WB_t + \beta_2 REV_{t-1} + \beta_3 RET_{t-1} + \beta_4 \Delta SIZE_{t+2} + \varepsilon_{t+2}$$

Dependent Variables[a]	Pred. Sign of *WB*	Independent Variables					Adj. R^2	n
		Intercept	*WB*	*REV*	*RET*	*ΔSIZE*		
ΔDIRECTORS	−	−0.14	−1.29	0.00	0.59	−0.00	4.2%	70
		(0.738)	(0.025)**	(0.527)	(0.230)	(0.756)		
Δ%INS_DIR	−	0.05	−0.09	−0.00	0.03	0.00	30.8%	70
		(0.003)***	(<0.001)***	(0.283)	(0.078)*	(0.829)		
Δ%BUSY	−	0.02	−0.07	0.00	−0.04	−0.00	−2.4%	70
		(0.565)	(0.091)*	(0.925)	(0.399)	(0.807)		
Δ%ATTEND	+	0.01	0.02	−0.00	0.01	0.00	−1.0%	66
		(0.368)	(0.111)	(0.889)	(0.122)	(0.878)		
Δ%CEO=COB%[b]	−	NA	0.01	0.00	0.16	−0.00	0.2%	72
			(0.506)	(0.831)	(0.745)	(0.820)		
ΔNEW_CEO%[c]	+	−1.19	1.34	0.00	0.80	−0.00	17.0%	68
		(0.005)***	(0.013)**	(0.154)	(0.100)	(0.121)		

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

p-values are reported in parentheses. p-values are one-sided for variables with directional predictions. We report (1 − p) values for coefficients that assume a sign opposite to the one predicted.

[a] We compare the value of each governance variable in year −1 (pre-allegation) to its value in year +2 (post-allegation).

[b] We report the results of an ordinal logistic regression, including the pseudo-R^2.

[c] We report the results of a logistic regression, including the pseudo-R^2.

Definition of variables (where firm subscripts are suppressed):

$\Delta DIRECTORS_{t+2}$ = number of directors in year +2 minus the number of directors in year −1;
$\Delta \%INS\_DIR_{t+2}$ = percent of inside directors in year +2 minus the percent of inside directors in year −1;
$\Delta \%BUSY_{t+2}$ = percent of "busy" directors in year +2 minus the percent of busy directors in year −1;
$\Delta \%ATTEND_{t+2}$ = percent board meeting attendance in year +2 minus the percent board meeting attendance in year −1;
$\Delta CEO = COB\%_{t+2}$ = an ordinal variable coded as 1 if the CEO was not the COB in year +2 but was in year −1, coded as −1 if the CEO was the COB in year +2 but was not in year −1, and coded as 0 otherwise;
$\Delta NEW\_CEO\%_{t+2}$ = an indicator variable coded as 1 if the firm replaced its CEO in year 0, year +1, or year +2, and 0 otherwise;
WB_t = 1 if the firm experienced a whistle-blowing event, 0 otherwise;
REV_{t-1} = total sales revenue (Compustat annual #12) for the year before the whistle-blowing event;
RET_{t-1} = market-adjusted buy-and-hold returns for 12 months, beginning with month −12 and ending with month −1, where month 0 is the event month; and
$\Delta SIZE_{t+2}$ = total assets in year +2 minus total assets in year −1.

to date, our focus is on (1) characteristics of firms targeted by employee whistle-blowers, (2) the economic consequences of such whistle-blowing revelations, and (3) firm responses to such allegations via subsequent governance changes.

Specifically, we report new empirical evidence on 218 employee whistle-blowing events between 1989 and 2004. Our paper adds to the literature on corporate governance in four ways. First, we conduct a cross-sectional analysis of the characteristics of firms subject to whistle-

blowing. Our evidence is broadly consistent with whistle-blowers targeting firms that were large, growing, successful, and highly regarded. Whistle-blowing tended to be more prevalent in firms that had downsized and in firms with worse governance. Finally, and not surprisingly, when the potential benefits to the whistle-blower increased (via the potential to share in any fraud-related settlement), employees were more likely to make public allegations of wrongdoing.

Second, we present evidence that, on average, whistle-blowing allegations had an immediate negative economic consequence for target firms. The average market-adjusted return in the five days around the day the allegation became public was −2.8 percent. Further, the stock market reaction was substantially more negative when the whistle-blower alleged earnings management (−7.3 percent).

Third, we find that whistle-blowing targets had more earnings restatements (Press sample), more shareholder lawsuits (Press sample), and relatively poor operating and stock return performance (press and OSHA samples) compared to a sample of propensity-score matched firms with similar characteristics. Whistle-blower allegations appear to be an early indicator of future negative economic consequences for targeted firms.

Fourth, we find that, on average, whistle-blowing targets exposed in the press improved several dimensions of governance relative to the year before the whistle-blowing event and relative to a matched sample of control firms. Interestingly, these governance improvements were not apparent for firms subject to whistle-blowing allegations that were not widely disseminated (OSHA sample). In sum, whistle-blowing is far from a trivial nuisance for firms targeted by allegations of corporate misconduct.

Our research represents early work on the role of financial whistle-blowing and provides indirect evidence on the effectiveness of Sarbanes-Oxley (SOX) Section 806 as a mechanism to expose agency problems at firms. We find that public whistle-blowing events are relatively infrequent, and a majority of the whistle-blowing allegations we identify are not directly related to the types of fraudulent behavior that precipitated the SOX legislation. Future research could examine (1) the impact of the whistle-blowing provisions in SOX on the incidence and types of whistle-blowing allegations and (2) whether investors, regulators, and other stakeholders benefit from (or overreact to) whistle-blowing allegations. Additionally, future research could examine how OSHA is performing in its role of adjudicating SOX-related whistle-blowing complaints. Finally, if one could obtain the data, an examination of the frequency, importance, and consequences of internal whistle-blowing would be a valuable contribution to the literature.

APPENDIX
VARIABLE DEFINITIONS

Capital Market Pressure

FIN_NEED = an indicator variable coded as 1 if the firm's free cash flow in the year prior to the whistle-blowing event is less than −0.1, and 0 otherwise. Similar to Dechow et al. (1996), we calculate free cash flow as cash flow less the average capital expenditure (Compustat annual #128) over the last three years, deflated by average total assets.

M&A = an indicator variable coded as 1 if the firm was involved in a merger or acquisition (Compustat AFTN #1) in year *t* or in any one of the three years prior to the whistle-blowing event, and 0 otherwise.

ABSDACC = the absolute value of discretionary accruals based on the modified Jones (1991) model for the fiscal year before the whistle-blowing event. Discretionary accruals are calculated as the fitted values from the following

Regression (2): $TOTACC = \beta_0 + \beta_1 \Delta REV + \beta_2 PPE + \beta_3 CFO + \beta_4 ROA$ + error. This regression is estimated each year for every two-digit SIC code conditional on having at least 10 firms in an industry. Note that *TOTACC* = total accruals, defined as income before extraordinary items minus cash flow from operations plus extraordinary items (CompUSA annual #18–Compustat annual #308 + Compustat annual #124), ΔREV = the change in revenue (Compustat annual #12), *PPE* = plant, property, and equipment (Compustat annual #7), *CFO* = cash from operations (Compustat annual #308), *ROA* = return on assets (Compustat #18). We scale each variable in Regression (2) by total assets (Compustat annual #6) at the beginning of the period.

CM_PRESSURE = the average scaled rank of the following variables: *FIN_NEED*, *M&A*, and *ABSDACC*. Each variable is coded such that higher values are consistent with greater capital market pressure. This measure of capital market pressure is averaged over the number of variables with non-missing data.

Growth

GROWTH = the scaled rank of sales growth, defined as the average growth in sales (Compustat annual #12) over the three years prior to the whistle-blowing event.

Past Performance

PAST_PERF = the scaled rank of market-adjusted buy-and-hold returns for 12 months, beginning from month −12 and ending month −1, where month 0 is the month the whistle-blowing was reported in the press or to OSHA.

Reputation

REPUTATION = an indicator variable coded as 1 if the firm has appeared on *Fortune*'s "Best Companies to Work For" list or on *Fortune*'s "Most Admired Companies" list in any of the five years prior to the whistle-blowing event, and coded as 0 otherwise.

Unclear Communication Channels

AGE = the number of years the firm has been listed on CRSP as of the year of the whistle-blowing event.

IND_CONC = the Herfindahl-Hirschman index based on the revenue for each of the firm's industry segments, calculated as the sum of squares of each industry segment's revenue as a percentage of total firm revenue. This variable is calculated for the year prior to the whistle-blowing event. See Bushman et al. (2004).

GEO_CONC = the Herfindahl-Hirschman index based on the revenue for each of the firm's geographic segments, calculated as the sum of squares of each

geographic segment's revenue as a percentage of total firm revenue. This variable is calculated for the year prior to the whistle-blowing event. See Bushman et al. (2004).

COMMUNICATION = the average scaled rank of the following variables: *AGE*, *IND_CONC*, and *GEO_CONC*. We reverse the sign of the rank of *AGE* such that lower values are consistent with worse internal communication channels. This measure of unclear communication channels is averaged over the number of variables with non-missing data.

Employee Downsizing

DOWNSIZE = the scaled rank of employee growth, defined as the average growth in the total employees (Compustat annual #29) over the three years prior to the whistle-blowing event.

Qui Tam

QUITAM = an indicator variable coded as 1 if the firm (1) is in the healthcare industry (2-digit SIC code = 80), or (2) appeared on the "100 Companies Receiving the Largest Dollar Volume of Prime Contract Awards" list in any of the three years prior to the whistle-blowing event, and 0 otherwise.

Firm Size

SIZE = the scaled rank of total sales revenue (Compustat annual #12) for the year before the whistle-blowing event.

Internal Control Weaknesses

ICW = the scaled rank of the fitted value from the following model, as estimated by Doyle et al. (2007): $ICW = \beta_2 + \beta_2 MARKETCAP + \beta_2 FIRM\_AGE + \beta_2 LOSSES + \beta_2 SEGMENTS + \beta_2 FOREIGN\_TRANS + \beta_2 EXTREME\_SG + \beta_2 RESTRUCTURE$ + error, where *MARKETCAP* = the log of the firm's market capitalization, *FIRM_AGE* = the log of the number of years the firm has CRSP data, *LOSSES* = an indicator variable equal to 1 if earnings before extraordinary items in the two most recent years sum to less than zero, and 0 otherwise, *SEGMENTS* = the log of the number of operating and geographic segments reported by the Compustat Segments database, *FOREIGN_TRANS* = an indicator variable equal to 1 if the firm has nonzero foreign translation, and 0 otherwise, *EXTREME_SG* = an indicator variable equal to 1 if year-over-year industry-adjusted sales growth falls into the top quintile, and 0 otherwise, and *RESTRUCTURE* = the aggregate restructuring charge in the two most recent years, scaled by the firm's market capitalization. *ICW* is measured as of the year immediately before the whistle-blowing event. We use coefficient values as reported in Doyle et al. (2007). Because of limited data availability, we set *SEGMENTS* equal to 0 if the necessary data is not available.

External Monitoring

BLOCK = percentage 5 percent block-holder ownership of the firm in the year prior to the whistle-blowing event obtained from Compact Disclosure.

PP = percentage pension ownership of the firm in the year prior to the whistle-blowing event obtained from Cremers et al. (2005).

EXT_MONITOR = the average scaled rank of the following variables: *BLOCK* and *PP*. Each variable is coded such that lower values are consistent with worse external monitoring. This measure of external governance is averaged over the number of variables with non-missing data.

Governance

EXECCOMP = the value of the CEO's in-the-money exercisable stock options, scaled by the CEO's wealth. CEO wealth is defined as the sum of the CEO's salary, bonus, the value of stock ownership, and the value of all in-the-money options. This variable is measured for the year prior to the whistle-blowing event.

DIRECTORS = the number of directors on the firm's board of directors in the year before the whistle-blowing event.

INSIDER_PCT = the percentage of directors who are insiders in the year before the whistle-blowing event.

CEO=COB = an indicator variable coded as 1 if the firm's CEO is also the chairman of the board (COB) in the year before the whistle-blowing event, and 0 otherwise.

BUSY_PCT = the percentage of directors who are "busy" in the year before the whistle-blowing event, where "busy" is defined as serving on at least three other corporate boards.

ATTEND = the percentage of board members who attended at least 75 percent of the board committee meetings in the year before the whistle-blowing event.

NEW_CEO = an indicator variable coded as 1 if the firm has a new CEO in the year before the whistle-blowing event, and 0 otherwise.

GOVERNANCE = the average scaled rank of the following variables: *EXECCOMP*, *DIRECTORS*, *INSIDER_PCT*, *ATTEND*, *CEO=COB*, *BUSY_PCT*, and *NEW_CEO*. Each variable is coded such that lower values are consistent with worse internal monitoring. This measure of internal governance is averaged over the number of variables with non-missing data.

REFERENCES

Anechiarico, F., and J. B. Jacobs. 1996. *The Pursuit of Absolute Integrity: How Corruption Control Makes Government Ineffective*. Chicago, IL: University of Chicago Press.

Armstrong, C., A. D. Jagolinzer, and D. F. Larcker. 2009. Chief executive officer equity incentives and accounting irregularities. *Journal of Accounting Research* 48 (2): 225–271.

Baucus, M., and J. Near. 1991. Can illegal corporate behavior be predicted? An event history analysis. *Academy of Management Journal* 34 (1): 9–36.

Beresford, D. R., N. deB. Katzenbach, and C. B. Rogers, Jr.. 2003. Report of investigation by the special investigative committee of the board of directors of WorldCom, Inc. Athens, GA.

Brickey, K. 2003. From Enron to WorldCom and beyond: Life and crime after Sarbanes-Oxley. Wash. U.L.Q: 357–401.

Bushman, R., Q. Chen, E. Engel, and A. Smith. 2004. Financial accounting information, organizational complexity and corporate governance systems. *Journal of Accounting and Economics* 37: 167–201.

Callahan, E. S., and T. M. Dworkin. 1994. Who blows the whistle to the media and why: Organizational characteristics of media whistleblowers. *American Business Law Journal* 32 (2): 151–184.

Cremers, K., J. Martijn, and V. B. Nair. 2005. Governance mechanisms and equity prices. *The Journal of Finance* 60 (6): 2859–2894.

Dechow, P., R. Sloan, and A. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13: 1–36.

Doyle, J., W. Ge, and S. McVay. 2007. Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics* 44: 193–223.

DuCharme, L. L., P. H. Malatesta, and S. E. Sefcik. 2004. Earnings management, stock issues, and shareholder lawsuits. *Journal of Financial Economics* 71: 27–49.

Dyck, I., A. Morse, and L. Zingales. 2008. Who blows the whistle on corporate fraud? Working paper, University of Toronto and The University of Chicago.

Economist. 2006. Tales from the back office. (March 23). Available at: http://www.economist.com/PrinterFriendly.cfm?story_id=5662830.

Francis, J. R., and C. S. Lennox. 2008. Selection models in accounting research. Working paper, University of Missouri and The Hong Kong University of Science and Technology.

General Accounting Office (GAO). 2002. *Financial Statement Restatements: Trends, Market Impacts, Regulatory Responses, and Remaining Challenges*. Report No. *03–138*. Washington, D.C.: General Accounting Office.

Gobert, J., and M. Punch. 2000. Whistleblowers, the public interest, and the public interest disclosure act 1998. *The Modern Law Review* 63: 25–54.

Government Accountability Office (GAO). 2006. *Financial Statement Restatements: Updates of Public Company Trends, Market Impacts, and Regulatory Enforcement Actions*. Report No. *06–678*. Washington, D.C.: Government Accountability Office.

Graham, J., C. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40: 3–73.

Hennes, K. M., A. J. Leone, and B. P. Miller. 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. *The Accounting Review* 83 (6): 1487–1519.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 28 (2): 193–228.

King, G. 1999. The implications of an organization's structure on whistleblowing. *Journal of Business Ethics* 20 (4): 315–326.

Luthans, B. C., and S. M. Sommer. 1999. The impact of downsizing on workplace attitudes. *Group & Organization Management* 24 (1): 46–70.

Miceli, M. P., and J. P. Near. 1985. Characteristics of organizational climate and perceived wrongdoing associated with whistle blowing decisions. *Personnel Psychology* 38: 525–544.

———, and ———. 1992. *Blowing the Whistle: The Organizational and Legal Implications for Companies and Employees*. New York, NY: Lexington Books.

———, and ———. 1994. Whistle blowing: Reaping the benefits. *The Academy of Management Executive* 8 (3): 65–73.

Miller, G. 2006. The press as a watchdog for accounting fraud. *Journal of Accounting Research* 44 (5): 1001–1033.

Moberly, R. E.. 2007. Unfulfilled expectations: An empirical analysis of why Sarbanes-Oxley whistleblowers rarely win. *William and Mary Law Review* 49: 65–155.

Near, J. P., and M. P. Miceli. 1996. Whistle-blowing: Myth and reality. *Journal of Management* 22: 507–526.

New York Times (*NYT*). 2009. *Dangerous Times*. (December 4): A34.

Public Company Accounting Oversight Board (PCAOB). 2004. *An Audit of Internal Control over Financial Reporting Performed in Conjunction with an Audit of Financial Statements*. Auditing Standard No. 2. Washington, D.C.: PCAOB.

Rothschild, J., and T. D. Miethe. 1999. Whistle-blower disclosures and management retaliation: The battle to control information about organization corruption. *Work and Occupations* 26 (1): 107–128.

Schmidt, M. 2003. *Whistle Blowing Regulation and Accounting Standards Enforcement in Germany and Europe—An Economic Perspective*. Berlin, Germany: Humboldt-Universität zu Berlin Wirtschaftswissenschaftliche Fakultät, Department of Business and Economics.

Securities and Exchange Commission (SEC). 2009. *Draft SEC Strategic Plan for 2010–2015*. Release No. 34–60799. Washington, D.C.: SEC.

Taxpayers Against Fraud (TAF). 2009. The False Claims Act Legal Center. Available at: http://www.taf.org/.

Zingales, L. 2004. Want to stop corporate fraud? Pay off those whistle-blowers. *AEI-Brookings Joint Center Policy Matters* (January 18).

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1273–1301

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1273

Auditors' Use of Brainstorming in the Consideration of Fraud: Reports from the Field

Joseph F. Brazel
North Carolina State University

Tina D. Carpenter
The University of Georgia

J. Gregory Jenkins
Virginia Polytechnic Institute and State University

ABSTRACT: Audit standards require auditors to conduct fraud brainstorming sessions on every audit. The Public Company Accounting Oversight Board has raised concerns about auditors' fraud judgments and the quality of their brainstorming sessions. We develop a measure of brainstorming quality to examine how it affects auditors' fraud decision-making processes. We test our measure using field survey data of auditors' actual brainstorming sessions for 179 audit engagements. Respondents report considerable variation in the quality of brainstorming in practice. We find some evidence that high-quality brainstorming improves the relations between fraud risk factors and fraud risk assessments. We also determine that brainstorming quality positively moderates the relations between fraud risk assessments and fraud-related testing. Our results suggest that the benefits of brainstorming do not apply uniformly, because low-quality sessions likely incur the costs of such interactions without receiving the attendant benefits. By documenting best practices from high-quality brainstorming sessions, our findings can inform auditors on how to improve their consideration of fraud.

Keywords: *brainstorming quality; fraud; fraud risk assessments; fraud risk response.*

We owe special thanks to Steve Kachelmeier (editor) and two anonymous reviewers. We are grateful for the helpful comments provided by Chris Agoglia, Ben Ayers, Ann Backof, Linda Bamber, Michael Bamber, Mark Beasley, Jodi Bellovary, Grey Burnett, Rex Yuxing Du, Brooke Elliott, Jenny Gaver, Steve Glover, Rick Hatfield, Rich Houston, Kathy Hurtt, Karla Johnstone, Bill Messier, Mitch Oler, Tom Omer, Linda Parsons, Mark Peecher, Denis Pelletier, Robin Pennington, Dave Piercey, Velina Popova, Bob Ramsay, Jane Reimers, Steve Salterio, Jaime Schmidt, Marjorie Shelley, Chad Simon, Jason Smith, CJ Song, Edward Swanson, Chris Wolfe, William Wright, Mark Zimbelman, and workshop participants at the University of Illinois, Texas A&M University, University of Nevada, Las Vegas, Baylor University, North Carolina State University, The University of Alabama, the 2007 AAA Annual Meeting, the North Carolina Office of the State Controller, First Citizens Bank, and the 2008 Auditing Section Midyear Meeting. We also thank Deloitte & Touche, Ernst & Young, Grant Thornton, KPMG, and PricewaterhouseCoopers for their support and participation in this project.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: July 2008
Accepted: December 2009
Published Online: June 2010

I. INTRODUCTION

Fraud detection is among the highest priorities for the accounting profession and standard-setters (Elliott 2002; PCAOB 2003, 2007, 2008). Statement on Auditing Standards (SAS) No. 99 (AICPA 2002), *Consideration of Fraud in a Financial Statement Audit,* requires brainstorming sessions on every audit as a means of improving auditors' fraud judgments. However, the Public Company Accounting Oversight Board (PCAOB) has expressed concerns about the actual conduct and quality of brainstorming sessions in practice (PCAOB 2007). Because these sessions are a relatively new requirement, little is known about how they occur in practice and how they influence fraud-related judgments. Accordingly, our study employs a field survey of actual audit engagements to pursue three research objectives. First, we study how audit teams are conducting brainstorming sessions in practice and develop a measure of brainstorming quality. This descriptive analysis should provide a basis for further modeling of brainstorming and future empirical research in team decision-making. Second, we investigate the comprehensive fraud decision-making process that audit teams are employing in practice, from the consideration of client fraud risk factors to their development of fraud risk assessments and fraud-related audit testing. Third, we examine how the quality of these brainstorming sessions influences auditors' consideration of fraud and whether variation in this quality has important consequences for the resulting fraud judgments.

This investigation is important for several reasons. First, prior psychology research suggests that the effectiveness of judgments and decisions likely depends on the quality of a team's interaction (e.g., Stasser 1999). We leverage the survey method to collect complex and rich data about brainstorming session quality and its underlying characteristics. Our measure of brainstorming session quality is multifaceted (i.e., 21-item scale) and is developed from a review of the psychology literature on brainstorming and team decision-making, the accounting literature on auditors' judgments and fraud, recent PCAOB inspection reports, and professional guidance. Our brainstorming session-quality measure consists of three overall elements: attendance and communication, structure and timing, and engagement team effort. To our knowledge, this is the first study in accounting or elsewhere to develop a measure of brainstorming quality and empirically examine its effects on team judgments.

Second, prior experimental studies have examined certain parts of the brainstorming process, but have not examined the entire decision-making process (e.g., Carpenter 2007; Hoffman and Zimbelman 2009). Moreover, the accounting literature suggests that the benefits derived from brainstorming depend on the brainstorming method used (Carpenter 2007; Hoffman and Zimbelman 2009; Lynch et al. 2009; Hunton and Gold 2010). Carpenter (2007) and Hoffman and Zimbelman (2009) document benefits to brainstorming versus not brainstorming. However, contemporary studies by Lynch et al. (2009) and Hunton and Gold (2010) suggest that the type of brainstorming is an important consideration for the benefits to be realized, as they find that some effective brainstorming methods are not predominant in practice.[1] While this prior research has

[1] Lynch et al. (2009) conclude that brainstorming effectiveness is higher for teams using group support systems relative to teams using traditional face-to-face brainstorming. In our sample, 80 percent of our observations used traditional face-to-face brainstorming. Hunton and Gold (2010) find that both the nominal group (where there is *no* communication among team members) and round robin brainstorming techniques are more effective than open brainstorming. However, in our sample, we find that 91 percent of our engagements employed open brainstorming. Thus, while these two studies provide means in which brainstorming can be improved, our descriptive data suggest that such changes would be significant and likely a substantial challenge for firms to undertake. Our finding of a preponderance of face-to-face communication and open brainstorming in practice likely suggests that unidentified institutional factors cause audit teams to employ these two potentially less effective methods (e.g., open brainstorming [versus round robin] allows a partner to better control the discussion).

provided some evidence on auditor fraud judgments in experimental settings, there is no empirical evidence on how audit teams are implementing brainstorming sessions in practice and whether the relations among their team judgments are improved by the *quality* of brainstorming sessions required by SAS No. 99. We contribute to the literature by using a field survey to collect auditors' client-specific brainstorming session data and the related fraud judgments. Because this methodology allows us to collect the underlying dynamics of audit teams' actual brainstorming sessions, and the resulting risk factors identified, fraud risk assessments, and related testing, we are able to examine multiple fraud decisions and relationships across a number of diverse audit engagements. Thus, we provide insight into the comprehensive fraud decision-making process that audit teams employ in practice.

Third, our investigation is important to auditors and standard-setters. PCAOB inspection reports indicate variation among audit teams' brainstorming sessions, and auditors' testing is often not responsive to their fraud risk assessments (PCAOB 2007). Our study empirically assesses the PCAOB's findings by documenting how auditors are implementing SAS No. 99 in practice and by investigating the consequences of brainstorming quality on the link between fraud risk assessments and subsequent testing. Thus, this study's findings can inform audit practice and standard-setters as they reconsider the professional guidance on auditors' consideration of fraud (PCAOB 2004, 2007).

Our theoretical model is based on the psychology and accounting literatures, PCAOB inspection reports, and professional guidance. Based on this model, we hypothesize that brainstorming quality positively moderates the relations between (1) fraud risk factors and related fraud risk assessments and (2) fraud risk assessments and subsequent audit testing. We test our hypotheses with data from 179 audit engagements for which highly experienced audit partners and managers from the Big 4 firms and an international firm completed a field survey.

Prior experimental research has typically found that when fraud risk factors are present, fraud risk assessments tend to be higher (e.g., Nieschwietz et al. 2000). Further, prior experimental research has documented that auditors have difficulty responding to fraud risk assessments by altering the nature of their testing (e.g., Zimbelman 1997; Glover et al. 2003). Actual audit engagements are rich in context and the fraud risk assessment is typically determined after engagement *team* deliberation. Such an environment is difficult to create in an experimental setting. Therefore, there is a need to complement experimental fraud research with research employing non-experimental methods. Our field survey results indicate that fraud risk factors identified for a diverse sample of clients are positively correlated with auditor fraud risk assessments. In addition, we find some evidence that brainstorming quality positively moderates the relations between fraud risk factors and fraud risk assessments. We also find that auditors' responses to fraud risk assessments are contingent on the quality of brainstorming sessions. Specifically, when audit teams engage in higher quality brainstorming sessions, fraud risk assessments are more positively associated with changes to the nature, staffing, timing, and extent of fraud-related audit procedures.

While several experimental studies have predicted a link between fraud risk and related audit testing, they have typically not found this link for the nature or type of procedures employed (e.g., Zimbelman 1997; Glover et al. 2003; Asare and Wright 2004). We complement this literature by documenting the important role brainstorming quality can play in improving the link between auditors' fraud risk assessments and their subsequent testing; including the choice of the *nature* of procedures. We consequently illustrate that important and tangible benefits can be derived from higher quality brainstorming. On the other hand, our analyses suggest that under-auditing can occur when brainstorming session quality is low. Thus, some engagement teams likely incur the cost of brainstorming without receiving the attendant benefits of such interactions. These new findings triangulate and extend prior studies by offering important insights on the costs and benefits of fraud brainstorming.

Legal restrictions prohibit the PCAOB from distributing descriptive evidence about brainstorming sessions in practice (PCAOB 2007). Thus, it is difficult for individuals outside the PCAOB to discern if the PCAOB's concerns about brainstorming quality are widespread or perhaps isolated to only a few audit engagements. Our study captures brainstorming data that is often not documented and thus not available for PCAOB inspection teams to review (e.g., participation levels of engagement team members). Our survey data validate the concerns raised by the PCAOB by documenting considerable variation in the reported quality of brainstorming sessions in practice. More importantly, we find evidence that there are consequential effects of conducting low- and high-quality sessions. We also analyze data related to the misappropriation of assets and fraudulent financial reporting that were detected by the audit teams in our sample (i.e., instances of actual fraud detection). In additional analyses, we examine audit team effectiveness by analyzing the links between *ex ante* fraud judgments and *ex post* fraud detection. Last, our data reveal best practices for how auditors can improve the quality of their brainstorming sessions (e.g., attendance of an IT audit specialist, occurrence of the session early in the planning process).

Section II describes the background and hypotheses development. Sections III and IV provide the methods and results of the study, respectively. Section V concludes the paper.

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

A Model of Auditors' Use of Brainstorming in the Consideration of Fraud

Based on SAS No. 99, as well as prior accounting and psychology research, we develop a model of auditors' use of brainstorming in the consideration of fraud. SAS No. 99 provides guidance to improve the likelihood that auditors will detect fraud using a multi-phase approach (AICPA 2002, 2003). First, auditors collect information related to the risk of material misstatement due to fraud (e.g., data on equity-based compensation). Using such information, auditors brainstorm to identify fraud risk factors (e.g., incentives), synthesize this information to develop a fraud risk assessment, and develop a response to the risk assessments such as altering the staffing of the engagement, or modifying the nature, timing, and extent of audit procedures. Brainstorming sessions are intended to aid auditors in linking fraud risk factors to risk assessments and, in turn, foster the development of appropriate audit responses. As such, the model depicted in Figure 1 indicates that brainstorming should influence both phases of the fraud decision-making process such that the relations among fraud risk factors, risk assessments, and responses are positively moderated by the quality of the brainstorming session.

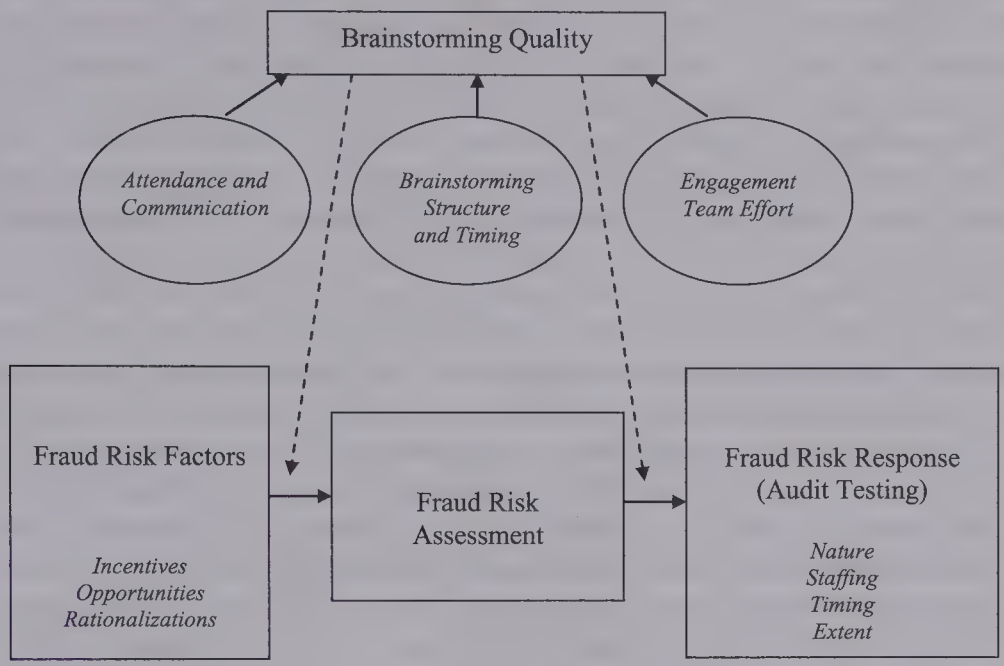
Fraud Risk Factors

Fraud risk factors are commonly categorized along three dimensions of the fraud triangle: incentives, opportunities, and rationalizations (AICPA 2002). Auditors synthesize their understanding of these risk factors with other information (e.g., preliminary analytical procedures) to develop a fraud risk assessment (AICPA 2002). Prior archival research on fraud risk factors documents relations between fraudulent financial reporting and financial statement elements, corporate governance quality, nonfinancial measures, and other observable data (e.g., Beasley 1996; Beneish 1997; McVay 2006; Brazel et al. 2009).

Fraud Risk Assessments

Experimental studies have examined the relationships between fraud risk factors and auditors' fraud risk assessments. While auditors' risk assessments are generally higher when fraud risk factors are present (Nieschwietz et al. 2000), more recent studies report that context has an important influence on auditors' judgments. Wilks and Zimbelman (2004) find that auditors who decompose fraud risk assessments are more sensitive to incentive and opportunity risks than auditors who provide holistic fraud risk assessments, but only in a low fraud risk context. Asare

FIGURE 1
Auditors' Use of Brainstorming in the Consideration of Fraud



This figure provides a model of audit teams' use of brainstorming in their consideration of fraud. This model is based on psychology and accounting research as well as the SAS No. 99 framework.

and Wright (2004) observe that, given a case where fraud was perpetrated, auditors who use a fraud risk checklist make less effective fraud risk assessments than those who use no checklist. Since the context of the experimental setting was important in these studies, and these studies were conducted before the implementation of SAS No. 99, it is important to re-examine the relations between fraud risk factors and fraud risk assessments using non-experimental methods. Further, none of these studies consider the possible influence of brainstorming quality on the links between risk factors and risk assessments.

Fraud Risk Responses

Although SAS No. 99 requires auditors to design audit procedures that are responsive to fraud risk assessments, research has not consistently found such a relationship. For example, Zimbelman (1997) illustrates limited evidence of an increase in the extent of audit procedures and no evidence of a change in the nature of procedures in response to fraud risk assessments. While Glover et al. (2003) find stronger evidence of an increase in the extent of audit procedures; they also fail to observe evidence of a change in the type of procedures employed in response to fraud risk. Additionally, Asare and Wright (2004) report no association between fraud risk assessments and planned audit procedures. However, these studies were conducted prior to SAS No. 99 and, thus, exclude the possibility that brainstorming could improve the sensitivity of fraud risk responses to fraud risk assessments.

Brainstorming Quality

Carpenter (2007) and Hoffman and Zimbelman (2009) experimentally study the influence of brainstorming sessions on the link between fraud risk and fraud risk assessment and the link between risk assessment and the nature of the audit procedures selected, respectively. Carpenter (2007) examines hierarchical teams that brainstorm and concludes that teams generate lists with more quality fraud risks than do individual auditors. Further, these teams provide more effective fraud risk assessments (i.e., higher when fraud was present than when it was not) than auditors working alone. Hoffman and Zimbelman (2009) investigate the role that brainstorming plays with regard to audit managers' planned audit procedures. They document that brainstorming (versus no brainstorming) leads to more effective modification of standard audit procedures in a specific high fraud risk case. While both of these studies provide important insights, neither captures the collective fraud process (i.e., the relations among fraud risk factors, risk assessments, and responses), nor do they examine actual audit teams where partners' judgments are likely very important.[2]

More importantly, contemporary experimental research by Lynch et al. (2009) and Hunton and Gold (2010) find that the brainstorming methods that are predominant in practice (e.g., face-to-face communication and open brainstorming) are less effective than alternative methods (e.g., electronic communication and round robin brainstorming). For a sample of audits where open brainstorming and face-to-face discussion were the norm, we examine if higher brainstorming quality can mitigate the negative effects documented by these contemporary studies. Furthermore, the PCAOB has expressed concern about brainstorming quality in practice (PCAOB 2007). However, because the PCAOB is legally restricted from distributing data from their inspections, it is difficult for those outside the PCAOB to discern if brainstorming quality issues are widespread or isolated, and the extent that session quality varies between engagements. Therefore, it is an open empirical question as to whether differences in brainstorming quality actually found in practice are large enough to lead to meaningful differences in audit teams' fraud risk assessments and audit testing.

There is a substantial body of psychology literature on the role of brainstorming. In general, this research reports mixed findings, with evidence suggesting that brainstorming can lead to both process gains and losses (e.g., Hill 1982; Diehl and Stroebe 1987). Most of these experimental studies were conducted with college students who held no task-specific expertise or knowledge and were asked to brainstorm about mundane problems (e.g., Diehl and Stroebe 1987). However, SAS No. 99 requires auditors with specific training and experience to conduct brainstorming sessions, a task with much at risk. Past studies of brainstorming also did not include hierarchical teams, eliminating the potential costs and benefits of normal hierarchies (cf., Hinsz et al. 1997). Hierarchical teams of auditors are likely to develop synergies by building on one another's ideas (e.g., Carpenter 2007). Conversely, the hierarchy of an audit team might exacerbate problems that can occur during brainstorming, such as evaluation apprehension, group think, and production blocking (e.g., Beasley and Jenkins 2003).

Psychology research has documented links between cooperative outcome interdependence, effectiveness of team communications, and resultant decisions. Under cooperative outcome interdependence, team members succeed or fail together, directly benefiting from each other's performance (De Dreu 2007). Audit teams have cooperative outcome interdependence because their

[2] Neither Carpenter (2007) nor Hoffman and Zimbelman (2009) include audit partners in their brainstorming teams. Given the authority level assigned to audit partners, it is likely that they wield substantial influence over fraud-related judgments. In our sample, partners attended every brainstorming session and were the largest contributors to the sessions (see Table 3). Also, whether the partner led the session and the amount of partner contribution were positively correlated with a number of fraud-related judgments (e.g., nature of testing; see Table 5). Our use of survey data from actual audit engagements allows us to include the role of partners.

team interaction in a brainstorming session depends on cooperation as opposed to competition. Cooperative outcome interdependence is related to the teams' ability to discuss shared and unshared information (Stasser 1999). Thus, the quality of information sharing or brainstorming among team members is likely to influence the effectiveness of the discussion and, in turn, the decisions of the team.

Building on professional guidance and the accounting and psychology literatures, our model in Figure 1 posits that the quality of the brainstorming session (1) positively moderates the relations between auditors' fraud judgments and (2) is based on three elements: attendance and communication, brainstorming structure and timing, and engagement team effort.

Attendance and Communication

Attendance and communication affect the quality of team interaction and discussion. Postmes et al. (2001) suggest that teams that promote critical thought are more likely to ensure careful evaluation of information and arrive at higher quality decisions. Schippers et al. (2003) find that the quality of team communication is enhanced when teams are more diverse and when they encourage open communication. Stasser (1999) illustrates that openness to ideas improves decision quality by encouraging the sharing of information. In line with this research, we expect that as more members of the engagement team attend and participate in the brainstorming session, there will be greater diversity and more sharing of information. This should improve the overall quality of the session and the responsiveness of fraud judgments.

Brainstorming Structure and Timing

The structure and timing of team discussions are important to the quality of team judgments. For example, increased time pressure exacerbates the tendency of teams to seek cognitive closure, reduces motivation, and increases reliance on heuristics (e.g., Kruglanski and Freund 1983; De Dreu 2003). Further, De Dreu (2003) suggests that pressure reduces team motivation to process information in a systematic and deliberate way, leading to reduced team effectiveness. Although time pressure is ever-present in the audit environment, auditors can potentially mitigate its negative effects by, for example, conducting brainstorming sessions early in the planning phase of the audit rather than later. Sessions held earlier in the planning process are expected to positively influence auditors' fraud judgments as the engagement team will have more time to implement the ideas endorsed during the session.

Engagement Team Effort

Engagement team effort is also an important determinant of the quality of teams' brainstorming sessions. There are two qualitatively different mechanisms of information processing: shallow, less critical, heuristic information processing and more argument-based, effortful, systematic information processing (Chaiken and Trope 1999). If the goal is to discover popular beliefs and preferences, team interaction tends to be preference-driven, leaving little opportunity for innovative decision making. In contrast, information-driven teams employ effortful systematic processing, communicating and integrating relevant information, to arrive at higher quality decisions (Stasser and Birchmeier 2003). Such an effortful process focus is evident in SAS No. 99-related guidance, which suggests that, for example, auditors individually identify risks and potential audit responses prior to brainstorming (AICPA 2003). These requirements should enhance each team member's involvement in the fraud audit process, augment their client-specific knowledge, and improve their contribution to the brainstorming session. Therefore, we expect that the more effort auditors exert prior to and during brainstorming, the higher the quality of team interaction and resultant fraud-related judgments.

In summary, higher quality brainstorming sessions are expected to improve (1) the links between fraud risk factors and related assessments and (2) the relations between fraud risk assessments and subsequent testing. That is, we expect session quality to positively moderate both relations, as stated in the following hypotheses:

H1: Fraud risk factors are more positively related to fraud risk assessments in high-quality brainstorming sessions than in low-quality brainstorming sessions.

H2: Fraud risk assessments are more positively related to fraud risk responses in high-quality brainstorming sessions than in low-quality brainstorming sessions.

III. METHOD

Participants

We approached contacts in the Big 4 firms and one other international firm to request participation in the study. Each firm agreed to participate. A secure Internet survey link was sent to each contact who distributed the survey to partners, directors, senior managers, and managers in their offices. A total of 367 auditors received the survey link, of which 179 completed the survey, resulting in a response rate of 48.8 percent. This rate exceeds the response rates in other auditing surveys (e.g., 16 percent response rate by Nelson et al. [2002]; 29.7 percent response rate by Dowling [2009]); however, it is lower than the exceptionally high 70.5 percent achieved by Gibbins et al. (2001).[3] Participants held positions of partner (56), director (2), senior manager (60), and manager (61). All 179 brainstorming sessions in our study reported to have an engagement partner in attendance.

Research Instrument

Similar to the method used by Gibbins et al. (2001), Nelson et al. (2002), and Dowling (2009), participants completed a field survey that required them to select one recently completed audit for which a fraud brainstorming session was held. They then responded to questions related to the consideration of fraud based upon their experiences on that engagement. Participants provided measures of fraud risk factors, their overall fraud risk assessment for the engagement, measures of audit responses to the fraud risk assessment, and objective and subjective data regarding the quality of the audit's fraud brainstorming session(s).

Fraud Risk Factors

Table 1 provides descriptive data related to fraud risk factors. Fraud risk factors elicited from participants were measured according to SAS No. 99 and prior research (e.g., Wilks and Zimbelman 2004) that characterize fraud risk factors as incentives, opportunities, or rationalizations. We measure incentives with seven questions related to market and debt incentives and opportunity with five measures of corporate governance quality, internal control strength, and accrual levels. We were unable to find research that suggests rationalization can be measured with multiple items, so we use one measure consistent with the wording in SAS No. 99. To facilitate comparisons across participants, we measure all but one of these items (i.e., public versus private company status) on a scale where 1 = extremely low and 10 = extremely high.

Factor analyses indicate that the items associated with market-related incentives and debt-related incentives load on separate factors. Because participants were allowed to provide data from

[3] Because not all individuals responded to our survey, we examine the potential for nonresponse bias. Oppenheim (1992) recommends comparing data from late respondents to early respondents as a way of assessing this bias. Accordingly, we compare the responses from the first quartile of respondents to those of the last quartile of respondents. There are no statistically significant differences between early and late responses on any of our hypothesized variables.

TABLE 1

Descriptive Statistics: Fraud Risk Factors

Item[a]	Response [n = 179] Mean	Std. Dev.
Market Incentive		
Percentage of publicly traded clients	67.03	(47.00)
Incentives or pressures for client to commit fraud	4.48	(2.14)
Pressure for client to meet earnings estimates	4.89	(2.57)
Extent of equity-based management compensation	4.28	(2.52)
Extent of budget-based management compensation	4.68	(2.40)
Debt Incentive		
Pressure for client to meet debt covenants	3.72	(2.68)
Amount of debt	4.93	(2.75)
Opportunity		
Opportunity for fraud to be perpetrated at client	4.90	(1.92)
Strength of corporate governance	6.85	(2.26)
Likelihood of management override of controls	4.32	(2.12)
Overall internal control effectiveness	6.74	(2.05)
Difference between net income and cash flow from operations	4.58	(2.34)
Rationalization		
Extent to which management is enabled to rationalize fraud	3.57	(2.07)

[a] With the exception of whether the client chosen was publicly traded (coded 1) or private (coded 0), all items are measured on a scale where 1 = extremely low and 10 = extremely high. Also, due to the positive wording of two opportunity items (strength of corporate governance and overall internal control effectiveness), responses for these two items are reverse-coded for data analysis.

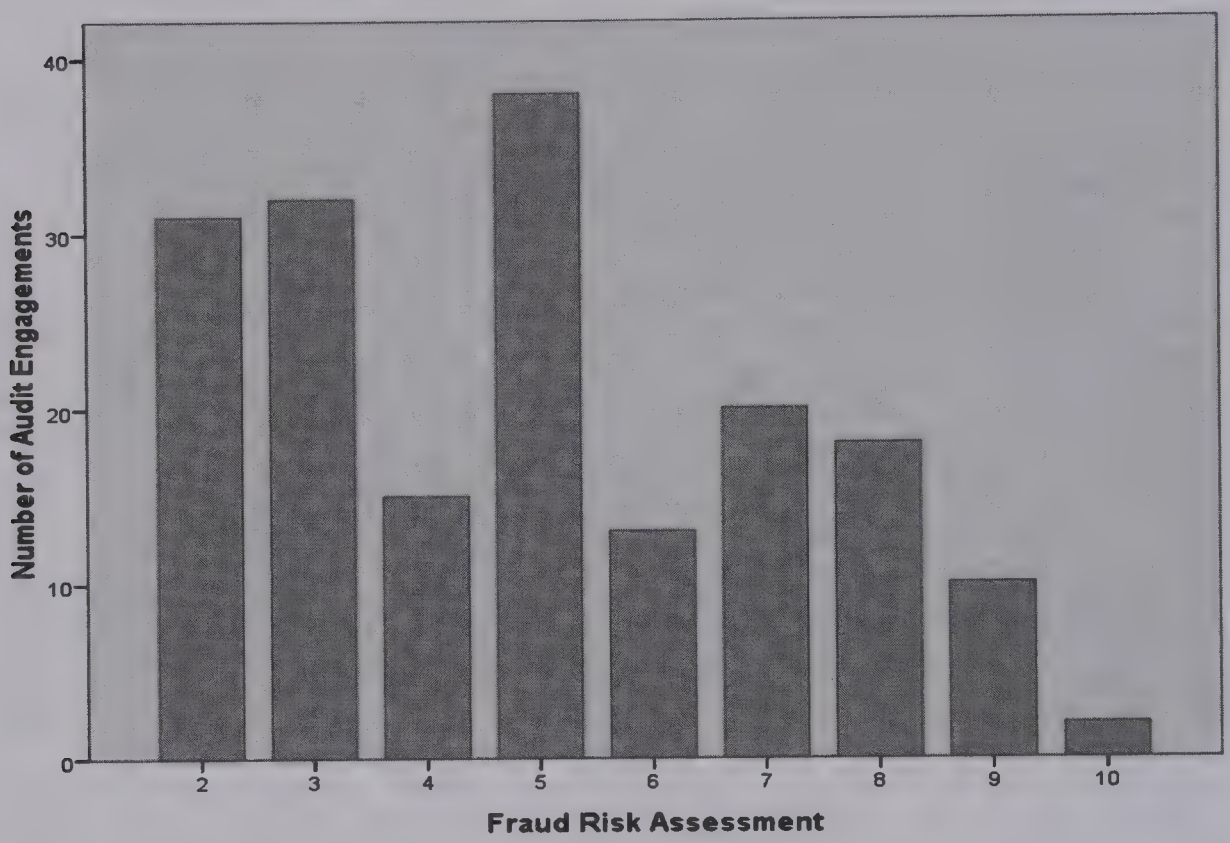
audits of both publicly traded and privately held companies, we decompose the incentive construct into "market incentive" and "debt incentive." We measure the four fraud risk factors as the mean of the participant's responses to the items classified under each factor.[4]

Fraud Risk Assessment and Fraud Risk Responses

Similar to Carpenter (2007), participants provided the overall fraud risk assessment for the engagement on a scale where 1 = extremely low and 10 = extremely high. The non-tabulated mean fraud risk assessment (standard deviation) for the sample was 4.92 (2.23). Figure 2 provides a frequency distribution of fraud risk assessments and illustrates substantial variation among the risk assessments, which ranged from 2 to 10 for our sample. After providing risk assessments, participants were asked to provide data on how the engagement team responded to the assessment by modifying the nature, staffing, timing, and extent of subsequent audit procedures. Table 2

[4] With only three exceptions, factor analyses of items combined to create fraud risk factors and fraud risk responses provide factor loadings in excess of 0.50 and Cronbach's alpha levels exceeding the generally accepted threshold of 0.70 (Nunnally 1978). For all combined measures, the percents of variances explained exceed 60.0 percent. With respect to the exceptions: (1) the factor loading for a measure of total accruals (0.283) was unacceptable and therefore excluded from subsequent analyses, (2) the remaining four measures of opportunity provide a moderate reliability level (Cronbach's alpha = 0.62) that is acceptable for testing relationships (e.g., Bamber et al. 1989), and (3) the Cronbach's alpha level for staffing (0.423) was low and we therefore provide results for staffing with the two staffing items combined and separated.

FIGURE 2
Frequency Distribution of Fraud Risk Assessments



This figure illustrates fraud risk assessments across the sample of actual audit engagements. Fraud risk assessment is measured on a scale ranging from 1 ("extremely low") to 10 ("extremely high").

provides descriptive statistics on the items used to measure fraud risk responses. The "nature" of responses was measured with two items: (1) the degree to which the number of testing procedures was increased, and (2) the degree to which the types of procedures were changed. "Staffing" was determined by both the degree to which more experienced staff or forensic specialists were used on the engagement. The "timing" of audit responses was measured by the degree to which the timing of audit procedures was changed on the engagement. The "extent" of auditors' responses was measured with three items: the extent to which (1) budgeted audit hours, (2) sample sizes, and (3) audit documentation review were increased. All eight items were measured on a scale where 1 = none and 10 = extensively. We measure the four fraud risk responses as the mean of the participant's answers to the underlying items.

Quality of the Fraud Brainstorming Session

We develop a 21-item measure to determine the quality of brainstorming sessions. Table 3 provides descriptive statistics for the 21 items used to measure quality based on three elements: attendance and communication, brainstorming structure and timing, and engagement team effort. Responses for each of the items were coded either 1 (high quality) or 0 (low quality), based upon the participant's response. Similar to DeFond et al.'s (2005) multiple measure of corporate gov-

TABLE 2

Descriptive Statistics: Fraud Risk Responses

Item[a]	Response [n = 179] Mean	Std. Dev.
Nature		
Increase the number of audit procedures	5.26	(2.43)
Change the types of audit procedures	5.56	(2.41)
Staffing		
Use more experienced auditors to perform audit procedures	4.36	(2.78)
Increase the use/consultation of forensic specialists	2.79	(2.73)
Timing		
Change the timing of audit procedures	3.83	(2.70)
Extent		
Increase budgeted hours	3.35	(2.43)
Increase sample sizes	3.73	(2.56)
Increase the extent of audit documentation review	4.71	(2.80)

[a] All items are measured on a scale where 1 = none and 10 = extensively.

ernance quality, we weigh each item equally to avoid adding additional subjectivity to the measure. Thus, we rate the total quality of a brainstorming session on a scale from 0–21. The nontabulated mean quality score for our sample was 10.56 (SD = 3.85), with a minimum score of 3 and a maximum score of 18. Figure 3 provides a frequency distribution of the brainstorming session quality scores, indicating substantial variation in brainstorming quality.

Attendance and Communication

Our measure includes 11 items related to which professionals led/attended the brainstorming sessions and the extent of their contribution. For example, proper tone at the top and a highly qualified person leading the session are important to the success of the brainstorming process (AICPA 2003; Landis et al. 2008). Further, whole engagement team participation allows more experienced auditors to share their fraud insights/experiences with less experienced team members (e.g., AICPA 2003). Therefore, if the brainstorming session is led by either the engagement partner or a fraud specialist, the engagement receives a coding of 1 (i.e., high quality) for this item. We measure the respondent's perceived level of contribution to the brainstorming session from staff, seniors, managers, partners, and various specialists, as well the level of the engagement management's openness to all ideas submitted during the brainstorming session. These items are measured on a scale from 1 (extremely low) to 10 (extremely high) and responses related to these items are coded high if they exceed the sample mean for the item (means provided in Table 3).

Brainstorming Structure and Timing

The survey contains four measures of brainstorming structure and timing. For example, practitioner guidance (e.g., AICPA 2003; Beasley and Jenkins 2003) suggests that the use of agendas should increase the effectiveness and efficiency of brainstorming sessions. Conversely, Asare and Wright (2004) find that the use of fraud checklists can impair auditor performance. Therefore, responses are coded as high quality if an agenda (checklist) is (is not) used.

TABLE 3

Descriptive Statistics: Quality of Fraud Brainstorming Session (21 Items)

Item	Response [n = 179] Mean	Std. Dev.
Attendance and Communication		
1. % of engagements where the session was led by the partner or forensic specialist	59.78	(49.22)
2. % of engagements where all levels of engagement team attended sessions	72.63	(44.50)
3. % of engagements where a fraud specialist attended the primary session	31.29	(46.34)
4. % of engagements where an IT audit specialist attended the primary session	69.27	(46.34)
5. % of engagements where a tax professional attended the primary session	62.56	(48.59)
6. Level of staff contribution to the session [n = 131][a,b]	4.32	(2.13)
7. Level of senior contribution to the session [n = 170][a,b]	6.03	(1.91)
8. Level of manager contribution to the session[a]	7.93	(1.48)
9. Level of partner contribution to the session[a]	8.41	(1.48)
10. Level of fraud specialist contribution to the session [n = 56][a,b]	7.01	(2.31)
11. Level of openness to ideas submitted during the session[a]	8.54	(1.42)
Brainstorming Structure and Timing		
12. % of engagements where an agenda was used during the session	83.80	(36.60)
13. % of engagements where a checklist was not used during the session	28.49	(45.34)
14. % of engagements where the primary session was held pre-planning or early in planning	64.81	(47.95)
15. % of engagements where a session was held at the end of the prior year's audit	16.20	(36.03)
Engagement Team Effort		
16. Time spent by the engagement team preparing for the session (in hours)	9.10	(15.51)
17. % of engagements that had more than one session	49.16	(50.13)
18. Total length of time for the sessions (in minutes)	93.68	(99.16)
19. % of engagements where team members were asked to identify risks prior to the session	67.60	(47.00)
20. Extent of discussion about how management might perpetrate fraud[a]	7.25	(1.69)
21. Extent of discussion about audit responses to fraud risk[a]	7.38	(1.60)

[a] Items are measured on a scale where 1 = extremely low and 10 = extremely high.

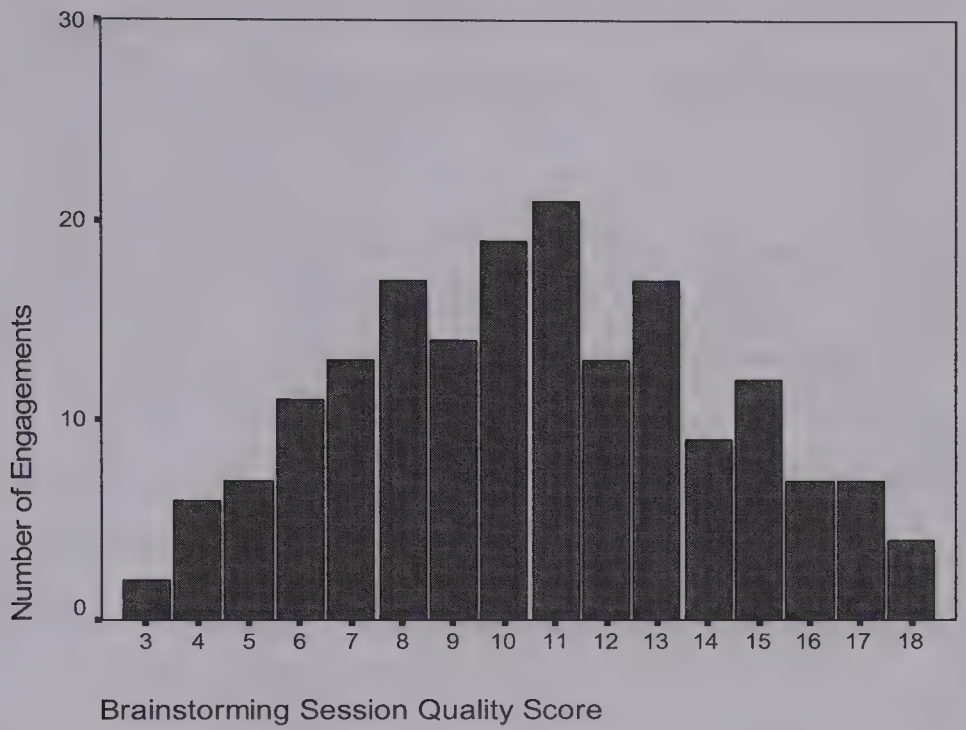
[b] Mean represents the average response in sessions where the professional(s) (e.g., staff, fraud specialist) was present (versus not present). Manager(s) and partner(s) were present for all observations.

Engagement Team Effort

We measure engagement team effort with six items. To illustrate, we deem responses as high if the time spent by the engagement team preparing for the brainstorming session exceeds the mean for our sample, if the engagement has more than one brainstorming session, or if the total length of time for the brainstorming session(s) exceeds the mean for our sample.[5]

[5] Factor analyses performed on the 21 categorical response items provide significant factor loadings ($p < 0.05$) for 20 of the 21 items on their respective constructs. Further, each of these 20 response items provide significant factor loadings ($p < 0.05$) on the overall construct of brainstorming quality. Both the Comparative Fit Index of 0.93 and the Tucker-Lewis Index of 0.93 are above 0.90, which is indicative of a good fit for our overall construct of brainstorming quality (Kline 2004). Only one item, percentage of engagements where all levels of the engagement team attended sessions, did

FIGURE 3
Frequency Distribution of Brainstorming Session Quality Scores



This figure illustrates brainstorming session quality scores across the sample of actual audit engagements. Brainstorming quality scores are formed using a 21-item measure with each item being equally weighted to determine a session's quality. Consequently, session quality scores can range from 0 to 21 points.

Control Variables

Table 4 provides descriptive statistics for our control variables. Prior research has found that client size, industry, and complexity impact auditor risk assessments and testing decisions (e.g., O'Keefe et al. 1994). We control for size by including a measure of the client's revenues in our analyses. We asked respondents to provide the client's industry and, similar to Johnstone and Bedard (2003), create dichotomous industry control variables to control for industry effects. As suggested by the audit fee and resource allocation literatures (e.g., Hackenbrack and Knechel 1997), we measure client complexity as the number of financial statement auditors assigned to the engagement divided by our size measure. Because we investigate fraud-related judgments, we control for whether fraud was detected on the engagement. We also control for the effects of other auditor or engagement team-related variables that could influence audit judgments (e.g., Bonner 1990). Specifically, with dichotomous firm control variables, we simultaneously control for the specific audit firm which performed the audit and the inclusion of one non-Big 4 firm in our sample. Also, our model includes the participant's general audit experience, position in the firm,

not load satisfactorily ($p = 0.70$). Omitting this item from the factor analyses and our session-quality measure in our regression analyses does not qualitatively change our results. We include this item in our analyses because the item was specifically identified as a deficiency by the PCAOB (PCAOB 2007).

TABLE 4

Descriptive Statistics: Control Variables

Item[a]	Response [n = 179] Mean	Std. Dev.
Client Size	3.06	(1.42)
Manufacturing industry	0.20	(0.40)
Retail industry	0.11	(0.31)
Government/not-for-profit industry	0.19	(0.39)
Energy industry	0.08	(0.28)
High tech/communications industry	0.14	(0.35)
Healthcare/pharmaceutical industry	0.07	(0.26)
Financial services industry	0.07	(0.25)
Miscellaneous industry	0.14	(0.35)
Client complexity	0.93	(0.65)
Fraud detect	0.24	(0.42)
Audit firm A	0.14	(0.35)
Audit firm B	0.19	(0.39)
Audit firm C	0.32	(0.47)
Audit firm D	0.16	(0.37)
Audit firm E	0.19	(0.39)
Audit experience	1.70	(0.86)
Position	2.29	(1.23)
Client experience	3.35	(2.05)
Industry experience	8.69	(6.73)
Engagement team expertise	7.50	(1.81)
Fraud training	1.42	(0.66)
Fraud experience	1.92	(0.71)
Restart	0.08	(0.28)
Time taken	22.29	(15.51)

[a] Items were measured as follows:

Client size is based on client revenues and coded as follows: 1 = < $100 million, 2 = $100 million −$500 million, 3 = > $500 million −$1 billion, 4 = > $1 billion −$5 billion, 5 = > $ 5 billion. Industry variables are dichotomous (1,0) industry variables for the following industries: manufacturing, retail, government/not-for-profit, energy, high tech/communications, healthcare/pharmaceuticals, financial services, and miscellaneous. Client complexity is measured as number of financial statement auditors assigned to the engagement (coded as follows: 1 = 0–5 auditors, 2 = 6–10 auditors, 3 = 11–15 auditors, 4 = 16–20 auditors, 5 = > 20 auditors)/client size. Fraud detect is coded 1 if the participant reported that fraud was detected at the client, 0 otherwise. Audit firm variables are dichotomous (1,0) firm variables for the four international and one national (i.e., non-Big 4) firm that participated in our study. Audit experience is coded as follows: 1 = 5–10 years, 2 = 11–15 years, 3 = > 15 years. Position is coded as follows: 1 = manager, 2 = senior manager, 3 = director, 4 = partner. Client experience is measured as the number of years the respondent had served the client. Industry experience is coded as follows: 1 = 0–5 years, 2 = 6–10 years, 3 = 11–15 years, 4 = 16–20 years, 5 = > 20 years. Engagement team expertise is measured on a scale where 1 = extremely low and 10 = extremely high. Fraud training is measured by the number of hours of fraud training the respondent had received in the last year and coded as follows: 1 = 0–8 hours, 2 = 9–16 hours, 3 = > 16 hours. Fraud experience is determined by the number of engagements the respondent served on which frauds were identified and coded as follows: 1 = 0, 2 = 1–2, 3 = >2. Restart is coded 1 if the participant stopped and restarted the survey, 0 otherwise. Time taken is the time taken by the 164 respondents who did not stop and restart the survey.

experience with the client and the client's industry, extent of fraud training, experience level with actual financial statement fraud, their perception of the entire engagement team's expertise level, whether the respondent stopped and started the survey, and time taken to complete the survey.

Regression Models

To test H1 and H2, we estimate the following models via ordinal logistical regression and examine the italicized interactions:

H1:

$$FR = \beta_0 + \beta_1 MI + \beta_2 DI + \beta_3 OPP + \beta_4 RAT + \beta_5 SESSION\ QUALITY + \beta_6 MI \times SESSION\ QUALITY + \beta_7 DI \times SESSION\ QUALITY + \beta_8 OPP \times SESSION\ QUALITY + \beta_9 RAT \times SESSION\ QUALITY + \beta_{10-32} \text{CONTROL VARIABLES} + \varepsilon.$$

H2:

$$NATURE = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26} \text{CONTROL VARIABLES} + \varepsilon.$$

$$STAFFING = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26} \text{CONTROL VARIABLES} + \varepsilon.$$

$$TIMING = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26} \text{CONTROL VARIABLES} + \varepsilon.$$

$$EXTENT = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26} \text{CONTROL VARIABLES} + \varepsilon.$$

where:

FR = Fraud Risk Assessment, measured on a scale where 1 = extremely low and 10 = extremely high;

MI, DI, OPP, RAT = Market Incentive, Debt Incentive, Opportunity, and Rationalization; see the "Fraud Risk Factors" section for information on measurement;

SESSION QUALITY = see the "Quality of the Fraud Brainstorming Session" section for information on measurement; this variable is coded 1 if session quality for the engagement exceeds the mean for the sample (10.56), 0 otherwise; the use of an indicator variable is described in the "Correlation Matrix and Multicollinearity" section below;

NATURE, STAFFING, TIMING, EXTENT = see the "Fraud Risk Assessment and Fraud Risk Responses" section for information on measurement;

CONTROL VARIABLES (*CLIENT SIZE, INDUSTRY VARIABLES, CLIENT COMPLEXITY, FRAUD DETECT, AUDIT FIRM VARIABLES, AUDIT EXPERIENCE, POSITION, CLIENT EXPERIENCE, INDUSTRY EXPERIENCE, ENGAGEMENT TEAM EXPERTISE, FRAUD TRAINING, FRAUD EXPERIENCE, RESTART, TIME TAKEN*)

= see the "Control Variables" section and Table 4 for information on measurement.

IV. RESULTS

Descriptive Statistics of Brainstorming Practices

Given our sample of 179 recently completed audit engagements across five public accounting firms, Table 3 provides descriptive evidence regarding how selected fraud brainstorming sessions are being performed in practice. First, our participants report that a partner or forensic specialist led the session only 60 percent of the time, pointing toward the important role audit managers and seniors probably play in determining the quality of the session. Second, consistent with recent PCAOB inspection team reports (PCAOB 2007), we find evidence that the attendance of specialists (fraud, tax, IT) in sessions varies. Third, our results are consistent with the hierarchical structure of engagement teams in that the level of contribution to the session increases with seniority. Fourth, consistent with PCAOB inspection team findings (PCAOB 2007), for 35 percent of engagements, sessions were held *after* the early stages of planning. Last, engagement team preparation for the sessions tended to be substantial, sessions lasted an average of 1.5 hours, and 49 percent of engagements conducted more than one session.

Correlation Matrix and Multicollinearity

Table 5 presents a correlation matrix. To present a parsimonious correlation matrix, brainstorming quality items and control variables are excluded from Table 5 if they are not significantly correlated ($p < 0.05$) with at least four of the hypothesized variables.

We test our hypotheses via interactions within five ordinal logistical regression models. These interactions are the cross product term of two independent variables (e.g., fraud risk $\times$ session quality). In such regressions, the interactive term is likely to covary with the main effects constituting the interaction term. These regressions often exhibit multicollinearity problems (Echambadi and Hess 2007). Likewise, when we interact our 21-item measure for session quality with risk factors (H1) and risk assessments (H2), diagnostics suggest multicollinearity exists between the interaction terms and the variables that constitute the interaction (e.g., variance inflation factors [VIFs] in excess of 10, condition indices in excess of 30 with more than 50 percent of the variance of coefficient estimates associated with these indices). Multicollinearity, similar to a small sample size, biases against finding statistically significant main effects and interactions in multivariate regression (e.g., Belsley et al. 1980; Hair et al. 1998; Wooldridge 2009).

Prior studies have mitigated multicollinearity effects by transforming their continuous or scaled moderating variable into an indicator variable (e.g., Core et al. 2002; Ali et al. 2008; Cardinaels 2008). Thus, we transform our 21-item measure of brainstorming quality into an indicator variable, coding each engagement's total session quality as high or 1 (low or 0) if the total score for the engagement exceeds (is less than) the mean for our sample (versus using the 21-item measure).[6] We use this indicator variable to examine the main and interactive effects of session quality in all of our tabulated and non-tabulated regression analyses. Tests of multicollinearity for our H1 regression analyses (with the session quality indicator variable) indicate that all but one of the independent variables for H1 provide satisfactory multicollinearity diagnostics (e.g., Belsley et al. 1980; Neter et al. 1996; Kennedy 2008). The sole exception is the VIF for the interaction between opportunity and session quality, which is 11.71. Using condition indices and

[6] Alternatively, we could have split the sample at the median session quality score (11.00) and excluded firms at the median from our analyses. However, we would then have eliminated 21 observations with a session quality of 11.00 from our analyses. Given our relatively small sample size (179), our large number of control variables (23), that the mean and median for our sample are approximately the same (10.56 and 11.00, respectively), and partitioning the sample at the mean can be performed without eliminating observations (i.e., no observations had a session quality score = 10.56), we use the mean session quality score to create the indicator variable for session quality.

TABLE 5

Correlation Matrix

		MI[a]	*DI*	*OPP*	*RAT*	*FR*	*NATURE*	*STAFFING*	*TIMING*	*EXTENT*
DI[a]	Correlation[b]	***0.22***								
OPP[a]	Correlation	0.03	***0.29***							
RAT[a]	Correlation	***0.22***	0.13	***0.42***						
FR[a]	Correlation	***0.24***	***0.22***	***0.32***	***0.31***					
NATURE[a]	Correlation	***0.23***	0.09	0.07	***0.17***	***0.39***				
STAFFING[a]	Correlation	***0.19***	0.13	0.06	***0.16***	***0.36***	***0.61***			
TIMING[a]	Correlation	***0.22***	0.10	0.13	***0.20***	***0.31***	***0.45***	***0.43***		
EXTENT[a]	Correlation	***0.16***	0.09	0.13	***0.20***	***0.46***	***0.69***	***0.64***	***0.50***	
PTR	Correlation	0.08	0.13	0.02	0.10	0.14	0.15	***0.25***	***0.19***	***0.18***
IT	Correlation	***0.18***	−0.07	***−0.29***	−0.06	0.00	0.09	***0.18***	***0.17***	0.10
EARLY	Correlation	0.00	0.08	−0.01	−0.05	***0.15***	***0.18***	***0.15***	***0.19***	***0.17***
FD	Correlation	***0.20***	−0.02	***−0.15***	0.10	***0.16***	***0.24***	***0.29***	0.11	***0.19***
RD	Correlation	***0.18***	−0.02	***−0.17***	0.08	0.13	***0.31***	***0.26***	***0.17***	***0.28***
MC	Correlation	0.13	−0.06	−0.12	−0.04	***0.17***	***0.25***	***0.24***	***0.15***	***0.24***
PC	Correlation	0.01	−0.05	***−0.16***	−0.03	0.14	***0.19***	***0.20***	***0.15***	0.14
SQ[a]	Correlation	***0.19***	0.00	***−0.16***	0.05	0.13	***0.30***	***0.40***	***0.28***	***0.30***
EXP	Correlation	***0.16***	−0.05	***−0.29***	−0.04	0.06	***0.22***	***0.21***	0.12	0.06
SIZE	Correlation	***0.27***	−0.01	***−0.18***	−0.12	−0.08	0.12	***0.26***	***0.14***	0.06
		PTR	***IT***	***EARLY***	***FD***	***RD***	***MC***	***PC***	***SQ***	***EXP***
IT	Correlation	***0.29***								
EARLY	Correlation	***0.23***	0.07							
FD	Correlation	0.07	0.06	0.03						
RD	Correlation	0.11	0.07	0.01	***0.74***					
MC	Correlation	−0.13	−0.03	0.03	***0.34***	***0.43***				
PC	Correlation	0.03	−0.04	0.06	***0.35***	***0.32***	***0.52***			
SQ[a]	Correlation	***0.36***	***0.37***	***0.29***	***0.51***	***0.48***	***0.45***	***0.42***		
EXP	Correlation	−0.02	***0.19***	−0.01	***0.24***	***0.20***	***0.18***	***0.16***	***0.22***	
SIZE	Correlation	0.11	***0.32***	−0.05	***0.17***	***0.17***	0.11	0.09	***0.31***	***0.32***

(continued on next page)

TABLE 5 (continued)

Correlations in bold italics are significantly correlated (p-value < 0.05).

[a] The "Method" section provides information on the measurement of market incentive (*MI*), debt incentive (*DI*), opportunity (*OPP*), rationalization (*RAT*), fraud risk assessment (*FR*), *NATURE*, *STAFFING*, *TIMING*, *EXTENT*, and brainstorming session quality (*SQ*). For illustrative purposes, the scaled (1–21) version of the *SQ* variable is presented in this table.

[b] Pearson correlation statistic.

Variable Definitions:

PTR = whether the session was led by the partner or forensic specialist (coded = 1, 0 otherwise);
IT = whether an IT audit specialist attended the primary session (coded = 1, 0 otherwise);
EARLY = whether the engagement's primary session was held pre-planning or early in planning (coded = 1, 0 otherwise);
FD = extent of discussion about how management might perpetrate fraud, measured on a scale where 1 = extremely low and 10 = extremely high;
RD = extent of discussion about audit responses to fraud risk, measured on a scale where 1 = extremely low and 10 = extremely high;
MC = level of manager contribution to the session, measured on a scale where 1 = extremely low and 10 = extremely high;
PC = level of partner contribution to the session, measured on a scale where 1 = extremely low and 10 = extremely high;
EXP = engagement team expertise, measured on a scale where 1 = extremely low and 10 = extremely high; and
SIZE = client size based on client revenues and coded as follows: 1 = < $100 million, 2 = $100 million–$500 million, 3 = > $500 million–$1 billion, 4 => $1 billion–$5 billion, 5 => $5 billion.

variances to assess multicollinearity (e.g., Belsley et al. 1980; Core et al. 2002), we find that only one of the 33 condition indices for H1 exceeds the standard threshold of 30, and over 50 percent of the variances for the intercept and *session quality* are associated with this index (Belsley et al. 1980; Kennedy 2008). This is likely the result of session quality being included in the model for H1 five times (one main effect and four interactions).[7] Thus, for testing H1, multicollinearity is likely to bias against finding significant main and interactive effects for *session quality,* and caution must be taken in interpreting null findings for this variable. For H2 analyses (with the session quality indicator variable), both VIFs and condition indices/variances indicate no problems with multicollinearity, as all measures are below acceptable thresholds (Neter et al. 1996; Belsley et al. 1980; Kennedy 2008).[8]

Hypotheses Testing

Fraud Risk Factors, Fraud Risk Assessments, and Brainstorming Session Quality

Hypothesis 1 posits that fraud risk factors are more positively related to fraud risk assessments in high-quality brainstorming sessions than in low-quality sessions. Thus, H1 is supported by a regression result that provides positive and significant interactions between fraud risk factors and session quality on fraud risk assessments. Table 6 provides the results of H1 testing. For parsimony, only control variables with p-values < 0.10 are included in Tables 6 and 7. While the main effects of market incentive, opportunity, and rationalization are all positive and significant (p-values < 0.05), their interactions with brainstorming session quality are not. Thus, while we find that three of the four risk factors are positively associated with fraud risk assessments, we do not observe that these relations are stronger under higher quality brainstorming.

However, diagnostics suggest that multicollinearity influences the main and interactive effects of session quality in the regression model for H1 (specifically related to its interaction with opportunity). As suggested by Bradshaw et al. (2004) and Covrig et al. (2007), we complement our analysis of H1 with two separate regressions with session quality as low and high (two subsamples). We partition the sample at the mean brainstorming quality score (10.56), and re-perform the statistical tests described above for each of the two groups (removing the brainstorming quality variable). If session quality positively moderates the effect of fraud risk factors on fraud risk assessments, then the main effects of fraud risk factors on fraud risk assessments should be more significant for the high-quality subsample. We find this to be the case for the fraud risk factors of debt incentive and opportunity. For the high-quality group, tests of H1 (not tabulated) indicate at least marginally significant and positive relations between debt incentive and opportunity on fraud risk assessments (p-values = 0.06 and 0.04, respectively). For the low-quality group, the relations between debt incentive and opportunity on fraud risk assessments are not significant. Thus, there is limited evidence supporting H1.[9]

---

[7] Given the significant correlations between our fraud risk factors (see Table 5) and that auditors *simultaneously* incorporate fraud risk factors into one fraud risk assessment, not including all four fraud risk factors in the same multivariate regression could yield spurious results.

[8] To ensure that our interaction results are stable across various specifications, we analyze the following models: (1) with both main and interactive effects for continuous hypothesized variables, (2) with both main and interactive effects for continuous hypothesized variables (mean-centered), (3) with only interactive effects for continuous hypothesized variables (no mean-centering), and (4) with both main and interactive effects for continuous hypothesized variables (no mean-centering, but session quality as an indicator variable (the tabulated model)). Our primary inferences are similar across all four models: the interactions for H1 are largely not supported and the interactions for H2 are largely supported. It should be noted that Echambadi and Hess (2007) conclude that mean-centering does not alleviate collinearity problems in multivariate regressions.

[9] Notwithstanding our expectation, we find only limited evidence that the quality of brainstorming sessions moderates the relations between risk factors and assessments. Anecdotal evidence from conversations with partners at the Big 4 firms suggests one possible explanation for our results. These partners report that engagement team management come to

TABLE 6

H1 Testing: Regression Results for Fraud Risk Assessment[a]

Independent Variable[b]	Estimated Coefficient	Wald-Statistic	p-value
MI	0.271	2.96	0.043
DI	0.095	1.07	0.150
OPP	0.317	3.13	0.039
RAT	0.249	4.35	0.019
SESSION QUALITY	0.409	0.11	0.739
MI × *SESSION QUALITY*	0.001	0.00	0.499
DI × *SESSION QUALITY*	0.018	0.02	0.443
OPP × *SESSION QUALITY*	0.086	0.14	0.356
RAT × *SESSION QUALITY*	−0.132	0.71	0.398
CLIENT SIZE	−0.252	3.10	0.078
ENGAGEMENT TEAM EXPERTISE	0.200	3.75	0.053
FRAUD TRAINING	0.649	7.23	0.007
RESTART	−1.580	3.51	0.061
Model Chi-square statistic = 79.60 (p-value = < .001)			
Pseudo R^2 = . 362			

[a] Fraud risk assessment (*FR*) is measured on a scale where 1 = extremely low and 10 = extremely high.

[b] The "Method" section provides information on the measurement of market incentive, debt incentive, opportunity, rationalization. *SESSION QUALITY* is coded 1 if the total session quality of the engagement (described in the "Method" section) exceeded the sample mean (10.56), 0 otherwise. *CLIENT SIZE* is based on client revenues and coded as follows: 1 =< $100 million, 2 = $100 million–$500 million, 3 => $500 million–$1 billion, 4 => $1 billion–$5 billion, 5 => $5 billion. *ENGAGEMENT TEAM EXPERTISE* is measured on a scale where 1 = extremely low and 10 = extremely high. *FRAUD TRAINING* is measured by the number of hours of fraud training the respondent had received in the last year and coded as follows: 1 = 0–8 hours, 2 = 9–16 hours, 3 = > 16 hours. *RESTART* is coded 1 if the participant stopped and restarted the survey, 0 otherwise.

Fraud Risk Assessments, Fraud Risk Responses, and Brainstorming Session Quality

Hypothesis 2 posits that fraud risk assessments are more positively related to fraud risk responses in high-quality brainstorming sessions than in low-quality sessions. Table 7 presents results related to H2. For nature, staffing, timing, and extent, the interaction terms for fraud risk assessment and session quality are all positive and at least marginally significant (p-values < 0.10, see Panels A–D), thus supporting H2. When brainstorming session quality is higher, audit teams are more likely to add or change testing procedures, employ more senior staffing/fraud specialists, change the timing of tests, and increase the extent of testing in response to higher fraud risk

brainstorming sessions with a lot of information about the likelihood of fraud based on prior years' work and client continuation discussions. Specifically, partners indicate that fraud risk assessments are often based on the presence or absence of specific risk factors such that assessments are effectively predetermined by the partners prior to the brainstorming sessions. In such instances, brainstorming serves primarily to inform team members about the known risk factors, to discuss other possible factors that might not have been pre-identified and, more importantly, to give them an opportunity to discuss how best to respond to fraud risk. How previous client experiences, client continuation discussions, and other factors affect the brainstorming process and fraud-related judgments is a question for future research.

TABLE 7

H2 Testing: Regression Results for Fraud Risk Responses[a]

Panel A: Nature

Independent Variable[b]	Estimated Coefficient	Wald-Statistic	p-value
FR	0.229	5.22	0.022
SESSION QUALITY	−0.148	0.05	0.832
FR × SESSION QUALITY	0.196	2.26	0.067
CLIENT SIZE	0.273	3.92	0.048
HEALTHCARE/PHARMACEUTICAL INDUSTRY	−1.903	7.92	0.005
CLIENT COMPLEXITY	0.643	5.68	0.017
FRAUD DETECTION	0.667	3.98	0.046
AUDIT FIRM A	−1.294	6.14	0.013
AUDIT FIRM C	−0.888	3.88	0.049
AUDIT FIRM E	−1.294	6.14	0.013
ENGAGEMENT TEAM EXPERTISE	0.162	3.38	0.066
INDUSTRY EXPERIENCE	0.055	3.42	0.064
Model Chi-square statistic = 87.35 (p-value < .001)			
Pseudo R^2 = .390			

Panel B: Staffing

Independent Variable	Estimated Coefficient	Wald-Statistic	p-value
FR	0.154	2.30	0.130
SESSION QUALITY	−0.130	0.03	0.854
FR × SESSION QUALITY	0.260	3.84	0.025
CLIENT SIZE	0.504	12.52	<0.001
CLIENT COMPLEXITY	0.599	5.00	0.025
FRAUD DETECTION	0.792	5.44	0.020
Model Chi-square statistic = 94.62 (p-value < .001)			
Pseudo R^2 = . 414			

Panel C: Timing

Independent Variable	Estimated Coefficient	Wald-Statistic	p-value
FR	0.101	0.95	0.330
SESSION QUALITY	−0.939	1.71	0.191
FR × SESSION QUALITY	0.328	5.96	0.008
MANUFACTURING INDUSTRY	1.097	4.39	0.036
AUDIT FIRM A	−1.29	5.13	0.023
AUDIT FIRM B	−1.90	13.37	<0.001
CLIENT EXPERIENCE	−0.137	3.19	0.074
FRAUD EXPERIENCE	−0.426	3.23	0.072
Model Chi-square statistic = 78.31 (p-value < .001)			
Pseudo R^2 = . 358			

(continued on next page)

Panel D: Extent

Independent Variable	Estimated Coefficient	Wald -statistic	p-value
FR	0.259	6.64	0.010
SESSION QUALITY	0.095	0.19	0.891
FR × SESSION QUALITY	0.199	2.32	0.064
CLIENT SIZE	0.234	2.92	0.088
RETAIL INDUSTRY	−1.163	3.52	0.061
FRAUD DETECTION	0.920	7.50	0.006
AUDIT FIRM B	−1.094	4.91	0.027
AUDIT FIRM C	−1.072	5.66	0.017
AUDIT FIRM E	−1.225	5.58	0.018
AUDIT EXPERIENCE	−0.763	4.59	0.032
CLIENT EXPERIENCE	−0.130	3.10	0.078

Model Chi-square statistic = 85.52 (p-value < .001)

Pseudo R^2 = . 383

[a] The "Method" section provides information on the measurement of the fraud risk responses of *NATURE*, *STAFFING*, *TIMING*, and *EXTENT*.

[b] *FR* is fraud risk assessment, measured on a scale where 1 = extremely low and 10 = extremely high. *SESSION QUALITY* is coded 1 if the total session quality of the engagement (described in the "Method" section) exceeds the sample mean (10.52), 0 otherwise. *CLIENT SIZE* is based on client revenues and coded as follows: 1 =< $100 million, 2 = $100 million–$500 million, 3 => $500 million–$1 billion, 4 => $1 billion–$5 billion, 5 => $5 billion. *HEALTHCARE/PHARMACEUTICAL INDUSTRY* is a dichotomous industry variable coded 1 if healthcare/ pharmaceutical industry client, 0 otherwise. *CLIENT COMPLEXITY* is measured as the number of financial statement auditors assigned to the engagement (coded as follows: 1 = 0–5 auditors, 2 = 6–10 auditors, 3 = 11–15 auditors, 4 = 16–20 auditors, 5 = > 20 auditors)/*CLIENT SIZE*. *FRAUD DETECTION* is coded 1 if the participant reported that fraud was detected at the client, 0 otherwise. *AUDIT FIRM* variables are dichotomous (1,0) firm variables for the four international and one national (i.e., non-Big 4) firm that participated in our study. *ENGAGEMENT TEAM EXPERTISE* is measured on a scale where 1 = extremely low and 10 = extremely high. *INDUSTRY EXPERIENCE* is coded as follows: 1 = 0–5 years, 2 = 6–10 years, 3 = 11–15 years, 4 = 16–20 years, 5 = > 20 years. *MANUFACTURING INDUSTRY* is a dichotomous industry variable coded 1 if manufacturing industry client, 0 otherwise. *CLIENT EXPERIENCE* is measured as the number of years the respondent had served the client. *FRAUD EXPERIENCE* is determined by the number of engagements the respondent served on which frauds were identified and coded as follows: 1 = 0, 2 = 1–2, 3 = > 2. *RETAIL INDUSTRY* is a dichotomous industry variable coded 1 if retail industry client, 0 otherwise. *AUDIT EXPERIENCE* is coded as follows: 1 = 5–10 years, 2 = 11–15 years, 3 = > 15 years.

assessments. Consistent with H2, a higher quality brainstorming session appears to substantially improve the risk-based strategy prescribed by SAS No. 99 (i.e., a positive relation between risk assessments and responses).[10]

To further illustrate that relations between fraud risk assessments and responses are stronger when session quality is high, we examine the response of *nature* in cases where risk assessment is "high" (i.e., 7–10 on our ten-point scale or the top quartile of risk assessments). The mean responses under high- and low-quality sessions are 8.19 and 5.37, respectively ($p < 0.01$), suggesting that under-auditing could be occurring when brainstorming quality is low.[11]

[10] As noted in footnote 4, the reliability statistic for staffing was below a satisfactory level. Therefore, we examine the interaction posited by H2 in relation to our two measures of staffing: increase use/consultation of fraud specialists and use of more experienced auditors. Non-tabulated results indicate that the interaction term for fraud risk assessment and brainstorming quality is positive and marginally significant (p-values < 0.10) for both measures.

[11] In order to provide results comparable to prior experimental research examining frauds at public companies (e.g., Carpenter 2007), we examine our hypotheses with only the publicly traded companies in our sample (sample size = 120). Non-tabulated regression results are qualitatively similar to those tabulated in this study.

Supplemental Discussion and Analyses

Specific Brainstorming Quality Items Driving Effects

While our 21-item measure of brainstorming quality is motivated by the literature and appears to improve the responsiveness of fraud judgments, it might not be practical for audit teams to adjust all 21 items at once to improve their brainstorming sessions. We therefore examine if a smaller number of items could be identified that would produce the same moderating effects as our 21-item measure. As the correlation matrix (Table 5) indicates, it appears that seven brainstorming items are significantly correlated with a number of the hypothesized variables in this study. These items are: (1) whether the session was led by the partner or forensic specialist, (2) whether an IT audit specialist attended the primary session, (3) whether the engagement's primary session was held pre-planning or early in planning, (4) the extent of discussion about how management might perpetrate fraud, (5) the extent of discussion about audit responses to fraud risk, (6) the level of manager contribution to the session, and (7) the level of partner contribution to the session. These are all controllable inputs that can be easily fostered by engagement management.

In testing H1 and H2, we replace the brainstorming quality indicator variable based upon 21 items with one based upon these seven items. Non-tabulated regression results reveal that this seven-item measure positively moderates the effect of rationalization on fraud risk assessments ($p = 0.03$). Recall that in Table 6, using a session-quality measure based on 21 items, all interactions between fraud risk factors and session quality are not significant. Similarly, the indicator variable for session quality based upon seven items positively moderates the relations between fraud risk assessments and the response factors of nature and staffing (p-values < 0.05). It is especially promising that concentrating on the aforementioned seven items could improve the often tenuous link between fraud risk assessments and the nature of testing. While our results using only the seven items do not provide exactly the same results as those tabulated with the 21-item measure, it appears that ensuring these seven items occur is a practical first step, or set of best practices, to improve the quality of brainstorming.[12]

Control Variables

With respect to the direct effects of control variables on fraud risk assessment in Table 6, we observe a negative relation between client size and fraud risk assessments. Consistent with Beasley et al. (1999), who find that fraud can be more prevalent at smaller companies; auditors tend to assess fraud risk higher for smaller clients. We find that engagement team expertise has a positive association with risk assessments, suggesting that auditors with more expertise are assigned to riskier engagements. We also find a positive effect for fraud training on risk assessment. Finally, participants who stopped and restarted the survey provided lower fraud risk assessments than other participants. Given this result, we re-perform our analyses of H1 and H2 without the participants who stopped and restarted and observe qualitatively similar results.[13]

The mean fraud risk assessment for our sample (4.92 on a ten-point scale) could be considered high, as the base rate for fraud is fairly low. On the other hand, given the relatively high rate of fraud detection in our sample (24 percent), a higher mean fraud risk assessment might be expected. In fact, the mean fraud risk assessment for the subset of audits where fraud was detected ($n = 43$) is a surprisingly low 5.44, and is only marginally higher than the mean for the non-fraud subset (mean = 4.75; $p = 0.09$; non-tabulated t-test). Moreover, fraud detection is not signifi-

[12] When we perform the same analyses with the remaining 14 items, there are no significant results for H1 or H2. This suggests that the seven items described above are driving our results.

[13] It is unclear why these participants provided lower fraud risk assessments. A review and analysis of the data for these participants reveal no clear explanation of the variable's significance.

cantly associated with fraud risk assessment in Table 6 (non-tabulated $p > 0.50$). Thus, while auditors in our sample were generally adept at incorporating fraud risk factors into fraud risk assessments, given *ex post* evidence of fraud detection, our results suggest that auditors still could improve the accuracy of their fraud risk assessments.

With respect to fraud responses, Table 7 illustrates that, as client size increases, auditors appear to be more likely to adjust the nature, staffing, and extent of testing. The larger audit fees/budgets for bigger clients typically allow for more testing when needed. Larger clients also tend to be more prestigious and, if needed to respond to fraud risks, engagement leaders can typically acquire additional audit professionals for the engagement. The negative relation between the healthcare/pharmaceutical industry and the nature of testing likely reflects that this industry was the only industry in our sample not to have fraud detected during the audit (possibly due to the industry being heavily regulated).

As one would expect, as client complexity increases, audit teams tend to alter the nature of testing and use more competent staff. Of particular interest is the positive link between the detection of fraud and the nature, staffing, and extent of procedures. These responses appear to be integral to detecting fraud, and it could be important for auditors to adjust their testing in multiple ways to improve fraud detection. It is also interesting to note that fraud detection is not significantly associated with the timing of tests. Consistent with this finding, we observe negative relations between both client experience and fraud experience and the timing of testing. These combined findings suggest that experienced auditors likely question the effectiveness of changing the timing of testing to detect fraud. However, this relation could be industry specific, as we find a positive association between the manufacturing industry and the timing of tests. Given the risks associated with manipulating ending inventory, auditors of manufacturing clients appear to be more apt to change the timing of their inventory examinations.

We find that audit firms are related to a number of our responses. Firms have different audit approaches or structures, and these differences can lead to different audit judgments (Prawitt 1995; Hyatt and Prawitt 2001). As one would expect, higher levels of engagement team expertise and industry experience are positively associated with the difficult task of changing the nature of testing. Last, we observe negative associations between both audit experience and client experience and the extent of the response. Given the positive relationship between fraud detection and the extent of testing (i.e., evidence that changing the extent of procedures can be an effective fraud detection tool), these combined findings suggest a negative effect of longer auditor tenure on audit quality. Such relations could, to some extent, be mitigated by the Sarbanes-Oxley Act of 2002, which mandates a five-year rotation for the lead and concurring audit partners on publicly traded clients (U.S. House of Representatives 2002).

Other Data Issues

As previously noted, fraud was detected on 24 percent of the engagements in our sample, a much higher rate than reported by previous studies (e.g., Loebbecke et al. 1989; Nieschwietz et al. 2000). This high rate of fraud could be due to the fact that fraud cases were more salient to participants when they selected an audit as the basis for responding to our survey. We asked for detailed information on the types and sizes of the frauds that were detected. However, few participants chose to provide this information. One respondent reported a fraud related to management's manipulation of payroll reports, while another reported a fraud related to payments to a fictitious vendor. In comparison to the other industries in our sample, both retail and governmental/not-for-profit entities had unusually high rates of fraud detection (36 percent and 32 percent, respectively). In retail, the misappropriation of retail inventory, whether material or not, could have contributed to its high rate. For governmental and not-for-profit entities, the incentive to fraudulently misstate financial statements is lower than in other industries. Thus, it is possible

that some frauds in our sample were related to misappropriations of assets (versus fraudulent financial reporting) and were likely immaterial from the auditor's perspective.

Given the high rate of fraud in our sample, we consider whether our primary results hinge on recalled audits that detected fraud. We examine this concern in two ways. First, we control for whether fraud was detected in all of our analyses. We find support for H2 even though the fraud detection control variable is significant in several of our analyses (see Table 7, Panels A, B, and D). Second, in non-tabulated analyses, we remove the 43 engagements where fraud was detected and find qualitatively similar results (i.e., H1 (H2) is not (is) supported).[14] Thus, our results do not appear to be an artifact of engagements for which fraud was detected.

We also examine whether our brainstorming data presented in Table 3 are significantly skewed by these engagements. Non-tabulated t-tests for each of the 21 items find that none of the means are significantly different between audits where fraud was or was not detected. Likewise, correlations between the fraud detection control variable and the 21-item measure and indicator versions of brainstorming session quality are not significant (p-values > 0.50). Last, the mean brainstorming quality score for the fraud sample (10.56) is actually equal to the mean for our entire sample (10.56) and is not significantly different than the mean for the non-fraud sample (10.57; $p > 0.50$). Thus, it appears that the descriptive brainstorming data presented in Table 3 is not significantly skewed by our fraud observations and that, for audits where fraud was detected, participants were not biased toward providing responses reflecting high-quality brainstorming sessions. In summary, these additional analyses suggest that our results are not primarily attributable to the audits in our sample where fraud was detected.

From an association perspective, it might be that audits with higher fraud risk assessments prompt auditors to conduct higher quality brainstorming sessions and carefully link fraud judgments as precautionary measures. Consequently, our results might not be reflective of audits with lower fraud risk assessments. In Table 5, we find that session quality is only marginally correlated with fraud risk assessments ($p = 0.09$) and, in a multivariate setting (Table 6), we find that fraud risk assessments and session quality are not significantly related ($p > 0.50$). Still, we re-perform our regressions with only audits where the fraud risk assessment ranged from 2–6 (removing engagements with the highest fraud risk assessments (7–10 or the top quartile of fraud risk assessments)). As noted in Figure 2, no respondents provided a fraud risk assessment of 1. Non-tabulated results related to this subset of less risky audits are qualitatively similar to those tabulated. Thus, our results do not appear to be driven by the riskier audit engagements in our sample.

V. CONCLUSION

SAS No. 99 requires brainstorming sessions on every audit (AICPA 2002), but the PCAOB has raised concerns about the quality of these sessions in practice (PCAOB 2007). This is the first study, to our knowledge, to consider whether and how the *quality* of brainstorming influences team decision-making.

Our study makes several contributions to research and practice. Leveraging the field survey method, we develop and test a measure of brainstorming quality using data collected from 179 actual audit engagements. We find that fraud risk factors are positively correlated with fraud risk assessments, and we find some evidence that brainstorming quality moderates these relations. More importantly, we find that when brainstorming quality is perceived to be higher, risk assessments are more positively related to the nature, staffing, timing, and extent of audit procedures. Achieving a high level of brainstorming quality appears to improve the audit team's consideration

[14] In tests of H1, excluding the 43 engagements where fraud was detected, all independent variables of interest are not significant (p-values > 0.50). In tests of H2, all independent variables of interest are significant (p-values < 0.10).

of fraud by cultivating a broader set of responses to identified fraud risks. However, we also observe that under-auditing can occur when brainstorming quality is low.

Prior and contemporary studies provide evidence that the extent that auditors benefit from brainstorming depends on the brainstorming method employed (Carpenter 2007; Hunton and Gold 2010). Archival research on fraud has investigated data from multiple public companies, but has not examined auditors' judgments in practice (e.g., Beasley 1996). Our results thus triangulate prior experimental and archival research and extend the fraud literature by providing new and important insights on the costs and benefits of fraud brainstorming for actual audit teams. Further, the descriptive data contained herein further enhance our understanding of the fraud decision-making and brainstorming processes. These data should provide a basis for further modeling of brainstorming and future empirical research in team decision-making.

While it might seem obvious (and arguably required by SAS No. 99) that high-quality brainstorming sessions should exist in practice, our results and recent PCAOB findings indicate substantial variation in brainstorming quality (PCAOB 2007). Indeed, auditors continue to seek better brainstorming methods (Landis et al. 2008). Thus, we contribute to practice by providing auditors and regulators with descriptive data that (1) complement previous PCAOB inspections and (2) highlight best practices for auditors to improve their fraud judgments (e.g., attendance of an IT audit specialist, and occurrence of the session early in planning).

Our study is subject to several limitations. Allowing participants to select their own audits could have biased their responses toward engagements that were memorable, such as when fraud was found. Our approach might have also led participants to select sessions they perceived as high quality or to report in such a way as to bolster their perception of having conducted a high-quality session. Our sample contains a fraud detection rate of 24 percent, which is likely higher than the actual rate of fraud in the population and is higher than the rate reported in earlier studies. Unfortunately, few participants reported details related to the frauds they identified, so we have little understanding of their nature. Sensitivity analyses to control for these engagements suggest that reported results are not driven by engagements on which fraud was detected or fraud risk was high. We also find that, for audits where fraud was detected, participants were not biased to report higher quality brainstorming sessions.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 2002. *Consideration of Fraud in a Financial Statement Audit*. Statement on Auditing Standards No. 99. New York, NY: AICPA.

———. 2003. *AICPA Practice Aid Series, Fraud Detection in a GAAS Audit*. SAS No. 99 Implementation Guide. New York, NY: AICPA.

Ali, A., S. Klasa, and O. Z. Li. 2008. Institutional stakeholdings and better-informed traders at earnings announcements. *Journal of Accounting and Economics* 46 (1): 47–61.

Asare, S. K., and A. M. Wright. 2004. The effectiveness of alternative risk assessment and program planning tools in a fraud setting. *Contemporary Accounting Research* 21 (2): 325–352.

Bamber, E. M., D. Snowball, and R. M. Tubbs. 1989. Audit structure and its relation to role conflict and role ambiguity: An empirical investigation. *The Accounting Review* 64 (2): 285–299.

Beasley, M. S. 1996. An empirical analysis of the relation between board of director composition and financial statement fraud. *The Accounting Review* 71 (4): 443–465.

———, J. V. Carcello, and D. R. Hermanson. 1999. *Fraudulent Financial Reporting: 1987–1997 An Analysis of U.S. Public Companies*. New York, NY: Committee of Sponsoring Organizations of the Treadway Commission.

———, and J. G. Jenkins. 2003. A primer for brainstorming fraud risks. *Journal of Accountancy* 196 (6): 32–38.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York, NY: John Wiley.

Beneish, M. 1997. Detecting GAAP violations: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16 (3): 271–309.

Bonner, S. E. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* 65 (1): 72–92.

Bradshaw, M. T., B. J. Bushee, and G. S. Miller. 2004. Accounting choice, home bias, and U.S. investment in non-U.S. firms. *Journal of Accounting Research* 42 (5): 795–841.

Brazel, J. F., K. Jones, and M. F. Zimbelman. 2009. Using nonfinancial measures to assess fraud risk. *Journal of Accounting Research* 47 (5): 1135–1166.

Cardinaels, E. 2008. The interplay between cost accounting knowledge and presentation formats in cost-based decision-making. *Accounting, Organizations and Society* 33 (6): 582–602.

Carpenter, T. D. 2007. Audit team brainstorming, fraud risk identification, and fraud risk assessment: Implications of SAS No. 99. *The Accounting Review* 82 (5): 1119–1140.

Chaiken, S., and Y. Trope. 1999. *Dual-Process Theories in Social Psychology*. New York, NY: Guilford Press.

Core, J. E., W. R. Guay, and S. P. Kothari. 2002. The economic dilution of employee stock options: Diluted EPS for valuation and financial reporting. *The Accounting Review* 77 (3): 627–652.

Covrig, V. M., M. L. DeFond, and M. Y. Hung. 2007. Home bias, foreign mutual fund holdings, and the voluntary adoption of international accounting standards. *Journal of Accounting Research* 45 (1): 41–70.

De Dreu, C. K. W. 2003. Time pressure and closing of the mind in negotiation. *Organizational Behavior and Human Decision Processes* 91 (2): 280–295.

——— 2007. Cooperative outcome interdependence, task reflexivity, and team effectiveness: A motivated information processing perspective. *The Journal of Applied Psychology* 92 (3): 628–638.

DeFond, M. L., R. N. Hann, and X. Hu. 2005. Does the market value financial expertise on audit committees of boards of directors? *Journal of Accounting Research* 43 (2): 153–193.

Diehl, M., and W. Stroebe. 1987. Productivity loss in brainstorming groups: Toward the solution of the riddle. *Journal of Personality and Social Psychology* 53 (3): 497–509.

Dowling, C. 2009. Appropriate audit support system use: The influence of auditor, audit team, and firm factors. *The Accounting Review* 84 (3): 771–810.

Echambadi, R., and J. D. Hess. 2007. Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science* 26 (3): 438–445.

Elliott, R. K. 2002. Twenty-first century assurance. *Auditing: A Journal of Practice & Theory* 21 (1): 139–146.

Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (3): 535–563.

Glover, S. M., D. F. Prawitt, J. J. Schultz, Jr., and M. F. Zimbelman. 2003. A test of changes in auditors' fraud-related planning judgments since the issuance of SAS No. 82. *Auditing: A Journal of Practice & Theory* 22 (2): 237–251.

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate Data Analysis*. 5th edition. Upper Saddle River, NJ: Prentice Hall.

Hackenbrack, K., and W. R. Knechel. 1997. Resource allocation decisions in audit engagements. *Contemporary Accounting Research* 14 (3): 481–499.

Hill, G. W. 1982. Group versus individual performance: Are N+1 heads better than one? *Psychological Bulletin* 91 (3): 517–539.

Hinsz, V. B., R. S. Tindale, and D. A. Volrath. 1997. The emerging conceptualization of groups as information processors. *Psychological Bulletin* 121 (1): 43–64.

Hoffman, V. B., and M. F. Zimbelman. 2009. Do strategic reasoning and brainstorming help auditors change their standard audit procedures in response to fraud risk? *The Accounting Review* 84 (3): 811–837.

Hunton, J., and A. Gold. 2010. A field experiment comparing the outcomes of three fraud brainstorming procedures: Nominal group, round robin and open discussion. *The Accounting Review* 85 (3): 911–935.

Hyatt, T. A., and D. F. Prawitt. 2001. Does congruence between audit structure and auditors' locus of control affect job performance? *The Accounting Review* 76 (2): 263–274.

Johnstone, K. M., and J. C. Bedard. 2003. Risk management in client acceptance decisions. *The Accounting Review* 78 (4): 1003–1025.

Kennedy, P. 2008. *A Guide to Econometrics*. 6th edition. Cambridge, MA: MIT Press.

Kline, R. B. 2004. *Principles and Practice of Structural Equation Modeling*. 2nd edition. New York, NY: Guilford Publications.

Kruglanski, A. W., and T. Freund. 1983. The freezing and unfreezing of lay inferences: Effects of impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology* 19 (5): 448–468.

Landis, M., S. I. Jerris, and M. Braswell. 2008. Better brainstorming. *Journal of Accountancy* 206 (4): 70–73.

Loebbecke, J. K., M. M. Eining, and J. J. Willingham. 1989. Auditors' experience with material irregularities: Frequency, nature, and detectability. *Auditing: A Journal of Practice & Theory* 9 (1): 1–28.

Lynch, A. L., U. S. Murthy, and T. J. Engle. 2009. Fraud brainstorming using computer-mediated communication: The effects of brainstorming technique and facilitation. *The Accounting Review* 84 (4): 1209–1232.

McVay, S. E. 2006. Earnings management using classification shifting: An examination of core and special items. *The Accounting Review* 81 (3): 501–531.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. *Applied Linear Statistical Models*. 4th edition. New York, NY: McGraw-Hill.

Nieschwietz, R. J., J. J. Schultz, and M. F. Zimbelman. 2000. Empirical research on external auditors' detection of financial statement fraud. *Journal of Accounting Literature* 19: 190–246.

Nunnally, J. D. 1978. *Psychometric Theory*. New York, NY: McGraw-Hill.

O'Keefe, T. B., D. A. Simunic, and M. T. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (2): 241–261.

Oppenheim, A. N. 1992. *Questionnaire Design, Interviewing, and Attitude Measurement*. New York, NY: Pinter Publishers.

Postmes, T., R. Spears, and S. Cihangir. 2001. Quality of decision making and group norms. *Journal of Personality and Social Psychology* 80 (6): 918–930.

Prawitt, D. F. 1995. Staffing assignments for judgment-oriented audit tasks: The effects of structured audit technology and environment. *The Accounting Review* 70 (3): 443–465.

Public Company Accounting Oversight Board (PCAOB). 2003. Professionalism is primary. Remarks delivered by Douglas R. Carmichael at AICPA National Conference, Washington, D.C., December 12.

———. 2004. PCAOB standing advisory group meeting: Meeting agenda, September 8–9. Available at: http://pcaobus.org/News_and_Events/Events/2004/09-08-09.aspx.

———. 2007. Observations on auditors' implementation of PCAOB standards relating to auditors' responsibilities with respect to fraud. Available at: http://pcaob.org/inspections/other/01-22_release_2007-001.pdf.

———. 2008. Report on the PCAOB's 2004, 2005, 2006, and 2007 inspections of domestic annually inspected firms. Available at: http://pcaobus.org/Inspections/Documents/2008_12-05_Release_2008-008.pdf.

Schippers, M. C., D. N. Den Hartog, P. L. Koopman, and J. A. Wienk. 2003. Diversity and team outcomes: The moderating effects of outcome interdependence and group longevity and the mediating effect of reflexivity. *Journal of Organizational Behavior* 24 (6): 779–802.

Stasser, G. 1999. The uncertain role of unshared information in collective choice. In *Shared Cognition in Organizations: The Management of Knowledge*, edited by Thompson, L., J. M. Levine, and D. M. Messick, 49–69. Hillsdale, NJ: Erlbaum.

———, and Z. Birchmeier. 2003. Group creativity and collective choice. In *Group Creativity: Innovation through Collaboration*, edited by Paulus, P. B., and B. A. Nijstad, 85–109. New York, NY: Oxford University Press.

U.S. House of Representatives. 2002. Sarbanes-Oxley Act of 2002. Public Law 107–204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Wilks, T. J., and M. F. Zimbelman. 2004. Decomposition of fraud-risk assessments and auditors' sensitivity to fraud cues. *Contemporary Accounting Research* 21 (3): 719–745.

Wooldridge, J. M. 2009. *Introductory Econometrics: A Modern Approach*. 4th edition. Mason, OH: South-Western.

Zimbelman, M. F. 1997. The effects of SAS No. 82 on auditors' attention to fraud risk factors and audit planning decisions. *Journal of Accounting Research* 35 (Supplement): 75–97.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1303–1323

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1303

Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items

Yun Fan
The University of Oklahoma

Abhijit Barua
Florida International University

William M. Cready
The University of Texas at Dallas

Wayne B. Thomas
The University of Oklahoma

ABSTRACT: McVay (2006) concludes that managers opportunistically shift core expenses to special items to inflate current core earnings, resulting in a positive relation between unexpected core earnings and income-decreasing special items. However, she further notes that this relation disappears when contemporaneous accruals are dropped from the core earnings expectations model. McVay (2006) calls for research to improve the core earnings expectations model and to provide additional cross-sectional tests of classification shifting. Using a core earnings expectations model that is not dependent on accrual special items, we show that classification shifting is more likely in the fourth quarter than in interim quarters. We also find more evidence of classification shifting when the ability of managers to manipulate accruals appears to be constrained and in meeting a range of earnings benchmarks. Overall, our evidence provides broad support for McVay's (2006) conclusion that managers engage in classification shifting. Our study also sheds new understanding of the conditions under which managers are more likely to employ classification shifting.

Keywords: *classification shifting; special items; earnings thresholds; accruals manipulation constraints.*

Data Availability: *Data are available from public sources identified in the paper.*

We are grateful for comments received from two anonymous reviewers, Sarah McVay, Mark Vargas, Han Yi, and workshop participants at The University of Oklahoma and The University of Texas at Dallas.

Editor's note: Accepted by Sanjay Kallapur.

Submitted: November 2008
Accepted: October 2009
Published Online: June 2010

I. INTRODUCTION

McVay (2006) hypothesizes that managers have incentives to report core expenses (defined as cost of goods sold and selling, general, and administrative expenses) as income-decreasing special items in an attempt to inflate core profitability. Managers are motivated to manage earnings in this way because market participants may be particularly interested in core earnings (or pro forma earnings) rather than bottom-line GAAP earnings (Bradshaw and Sloan 2002; Gu and Chen 2004; Kinney and Trezevant 1997), and core earnings typically receive higher valuation multiples than non-core earnings (Lipe 1986). By shifting core expenses to special items, the firm increases core earnings, while bottom-line net income remains unaffected.[1]

McVay (2006) presents empirical evidence supportive of classification shifting in the form of a *positive* relation between income-decreasing special items and unexpected core annual earnings. As income-decreasing special items increase, the firm tends to report higher than expected core earnings. However, McVay (2006, Table 10) also reports that the positive relation disappears (and, in fact, becomes *negative*) when contemporaneous accruals are dropped from the core earnings expectations model. She further cautions that the initial evidence in favor of classification shifting may be attributable to the use of contemporaneous accruals in the core earnings expectations model. More specifically, the core earnings expectations model (which includes accruals as a control for performance) could induce a mechanical relation between unexpected core earnings and special items (which are mainly accrual-based). Thus, there is currently some uncertainty regarding the extent to which income-decreasing special items include shifted core expenses.[2]

Research in the area of classification shifting is relatively new, and it is essential that subsequent research provides improvements to the core earnings expectations model. We first extend McVay's (2006) core earnings expectations model and show that as further controls for performance are included, the evidence in favor of classification shifting increases. We also present a series of empirical examinations of unexpected core earnings at the quarterly level that are robust to the underlying bias discussed in McVay (2006). As research in this area grows and researchers develop fundamental views, it is important that findings from initial studies are validated. Our tests provide broad support for McVay's conclusion of classification shifting by managers.[3]

A central aspect of our analysis is the use of quarterly rather than annual data. This distinction allows us to use the differential incentives to manipulate fourth versus interim quarter earnings as a platform for constructing tests of classification shifting. These examinations are inherently tests of a joint hypothesis—that managers both engage in classification shifting and are more prone to do so in fourth quarters. Consequently, our results also contribute to the literature on differential incentives/constraints facing managers attempting to manipulate fourth quarter earnings. Using quarterly data, we demonstrate that classification shifting is more pronounced in the fourth quarter than in interim quarters.

We next consider that classification shifting is one means of earnings manipulation, while others exist (e.g., accrual manipulation). To the extent that managers are more constrained in their ability to manipulate accruals (Brown and Pinello 2007), they may use classification shifting as an

---

[1] Unlike income classification shifting, accrual manipulation and real activities management affect reported bottom-line GAAP earnings.

[2] McVay (2006) provides several approaches to test for classification shifting, some of which are not necessarily expected to be affected by the mechanical relation between unexpected core earnings and (accrual) special items. She finds (1) subsequent reversal of shifted earnings is more likely when there is no opportunity to shift in the subsequent year, (2) shifting is more likely when firms just meet the analyst forecast, especially for high-growth firms, (3) shifting is more likely when the type of special item is easier to shift, and (4) abnormal returns in the subsequent year tend to be negatively related to shifted earnings as these earnings recur in core earnings, disappointing investors.

[3] In this study, we limit our tests to the relation between unexpected core earnings and special items in the *current* quarter. We do not examine the subsequent reversal of unusually high core earnings associated with special items.

alternative to inflate core earnings. Finally, we examine classification shifting to meet earnings thresholds such as analysts' forecasts, one-year-ago same-quarter earnings, and zero earnings. We do so, however, using a core earnings expectations model that does not rely on contemporaneous accruals.

In the next section, we discuss the background literature and analysis of the classification shifting issue identified in McVay (2006). In Section III, we develop our hypotheses. Section IV details the data, sample, and descriptive statistics. Section V describes the research design and reports tests of hypotheses, while Section VI concludes the study.

II. BACKGROUND LITERATURE AND ANALYSIS

Earnings management has been the subject of considerable academic research. Evidence of managers engaging in earnings management through accrual manipulation has been shown in many different contexts, for many different accruals, and in response to many managerial incentives, as reviewed by Healy and Wahlen (1999). A second channel through which earnings could be manipulated is real activities management, such as providing temporary price discounts to increase sales, cutting discretionary expenditures such as research and development and advertising, and overproducing inventory to reduce cost of goods sold (e.g., Baber et al. 1991; Bushee 1998; Gunny 2005; Roychowdhury 2006; Cohen and Zarowin 2008). Unlike accrual manipulation, which involves potential accounting fraud that brings about litigation risk to the firm and/or employment risk to the manager, real activities manipulation does not violate GAAP but does sacrifice firms' future economic benefits.

McVay (2006) investigates a third form of earnings management. She tests whether managers manipulate *core* earnings by reclassifiying ordinary operating expenses as special items. To the extent that investors fixate on core earnings instead of bottom-line GAAP earnings, managers may employ classification shifting to fool financial statement users by manipulating earnings numbers within the income statement. This is because financial statement users appear to recognize and distinguish the closeness to sales of individual line items and weigh them differently (e.g., Lipe 1986; Fairfield et al. 1996; Francis et al. 1996). Various studies document that analysts and investors pay more attention to "street" or pro forma earnings as defined by managers and view them as more value-relevant (Bradshaw and Sloan 2002; Bhattacharya et al. 2004; Gu and Chen 2004). Thus, it is likely that managers take advantage of the market's focus on core earnings instead of bottom-line GAAP earnings to misclassify some core expense items in the income statement. McVay (2006) finds that managers opportunistically shift core expenses to income-decreasing special items, and investors act as if they are surprised in the following period when these expenses shift back into core earnings in the following year.

For empirical tests of classification shifting, firms' earnings performance poses a challenge. This occurs because the magnitude and frequency of income-decreasing special items are markedly higher among firms experiencing poor performance (McVay 2006). Thus, there is a natural *negative* relation between earnings performance and the amount of income-decreasing special items. However, classification shifting predicts a *positive* relation between unexpected core earnings and the amount of income-decreasing special items (i.e., as core expenses are shifted to special items, unexpected core earnings and income-decreasing special items both increase).

The existing body of empirical evidence supportive of classification shifting centers around measures of expected core earnings, as developed in McVay (2006). A common feature of these models is the inclusion of contemporaneous accruals in the formation of expected core earnings values. This inclusion follows from the DeAngelo et al. (1994) linkage of accruals to underlying economic performance. As discussed by McVay (2006), conditioning core earnings expectations on contemporaneous accruals is problematic because this number includes accrual special items.

As these same accrual special item amounts also appear within the total special item amount, a situation is created wherein the dependent variable (i.e., unexpected core earnings) that is partially determined by variable *X* (accrual special items) is regressed on a variable (income-decreasing special items) that is also partially determined by variable *X*. This potentially creates a mechanical (positive) relation between the dependent variable and independent variable.

To illustrate, the basic McVay design employs a two-stage regression procedure. In the first stage, expected firm-year core earnings, $E(CE)$, is derived as a function of accruals in that year multiplied by an industry-determined multiplier β_a. The value of this multiplier averages 0.22 (see McVay 2006, Table 4). Consequently, a dollar of special item accrual expense/loss reduces $E(CE)$ by \$0.22 on average. The difference between actual core earnings and $E(CE)$ is identified as "unexpected core earnings" and is regressed on special items (SI) in a second stage analysis of the form:

$$[\text{Core Earnings}_{i,t} - E(CE_{i,t})] = \alpha_0 + \alpha_1 SI_{i,t} + \varepsilon_{i,t}. \quad (1)$$

A dollar of special item accrual expense does not impact core earnings (from which special items are excluded by construction) but, on average, reduces $E(CE_{i,t})$, and therefore increases $[\text{Core Earnings}_{i,t} - E(CE_{i,t})]$, by $\beta_a * SI_{i,t}$.[4] Definitionally, the special item accrual also increases $SI_{i,t}$, forcing $\alpha_1 > 0$ in cases where, in fact, no relation may exist between SI and true unexpected core earnings magnitudes.[5,6]

As McVay (2006) acknowledges (see her Table 10), when accruals are removed from the expectations model, the positive relation between unexpected core earnings and special items becomes negative. The flipped sign of the relation could be due to either insufficient control for performance or the possible bias described above. In the tests detailed in the next section, we exclude contemporaneous accruals to avoid introducing the potential bias and attempt to better control for the performance effect.

III. HYPOTHESES

We rely on comparative analyses of relative shifting activities between firm-groups that are *a priori* more versus less likely to engage in classification-shifting activities. Specifically, we use quarterly data, as opposed to annual data in McVay (2006), and partition the sample based on accrual manipulation constraints and earnings thresholds. To rule out the possibility of spurious accrual special item effects driving our results, we eliminate current-period accruals from the core earnings expectations model.[7] It is likely that the overall relation between unexpected core earnings and special items includes both classification shifting and a performance-driven effect. However, for certain firms and in certain quarters, classification shifting may be more prevalent because of greater managerial incentives or opportunities. Breaking down the data by specific quarters and by different levels of managerial incentives and opportunities gives us a better

[4] This statement holds so long as β_a is positive. Empirically, this assumption is true most of the time (81.5 percent in McVay [2006]).

[5] Gu and Chen (2004, footnote 10) recognize that the interrelationships between non-cash special items and accruals can create problems in research designs that include both variables. They show (in their Table 6, Panel B) that the persistence coefficient of special items changes sign after the inclusion of the accruals variable, which they interpret as the effect of noncash nonrecurring items being included in accruals.

[6] It is also straightforward to show that in a changes analysis (i.e., unexpected change in core earnings) negative accrual special items tend to increase expected change in core earnings and, as a consequence, decrease unexpected change in core earnings. This produces a mechanical negative relation between unexpected change in core earnings and income-decreasing special items.

[7] In addition, in a sensitivity test reported below, we specifically address whether the relation between our estimate of unexpected core earnings and income-decreasing special items is related to the amount of *contemporaneous* accrual special items. We find no evidence that it is.

understanding of whether and when classification shifting is more likely to occur.

Our motivation for using quarterly data comes from prior literature that suggests that earnings management activities are likely to differ between the fourth quarter and the interim quarters. Brown and Pinello (2007) specifically examine whether accruals are managed more in the fourth quarter relative to interim quarters. They contend that since the annual reporting process is subject to a more rigorous audit compared to interim quarters, it is more difficult for managers to manipulate accruals upward at the end of the year. Instead, managers are more likely to resort to expectations management in the fourth quarter as a way to meet analysts' forecasts. Their results confirm these views. However, Brown and Pinello (2007) conclude that the extent of expectations management in the fourth quarter seems too small to explain the ability to meet fourth quarter earnings expectations, given the decrease in accrual manipulation ability. Our test of classification shifting in the fourth quarter provides an additional explanation for their puzzling result. During the year, managers are more likely to manipulate accruals to meet analysts' expectations, but at the end of the year, managers' ability to manipulate accruals is constrained (and it is more difficult to manage expectations). This difficulty leads managers to resort to an alternative earnings management technique—classification shifting.

Consistent with the views expressed by Brown and Pinello (2007), Cornell and Landsman (1989) find that both price reactions and analysts' revisions are greater around earnings announcement in the fourth quarter than in interim quarters. They conclude that audited annual announcements are more informative. Prior research also provides evidence that the frequency and magnitudes of special items are significantly greater in the fourth quarter (Burgstahler et al. 2002; Kinney and Trezevant 1997).[8] We conjecture that these results indicate that not only do managers have more incentives to employ classification shifting in the fourth quarter, but they also have greater opportunity to camouflage core expenses with other (true) special items. This discussion leads to our first hypothesis:

H1: Managers shift core expenses to special items more in the fourth quarter.

We next consider that managers' propensity to shift core expenses in the current quarter may be affected by previous accruals manipulation activity. To the extent that managers have already managed accruals upward in previous quarters, the ability to further manage accruals upward in the current quarter is constrained. As a result, earnings management through income classification shifting may provide the better (or only) alternative. Barton and Simko (2002) measure firms' previous earnings management using the beginning level of net operating assets (scaled by sales).[9] They find that firms with higher net operating assets (i.e., more optimistic reporting in prior periods) are less likely to report earnings that just meet or beat analysts' forecasts (see Barton and Simko 2002, Table 5, Panel B). Related to our study, to the extent that accruals manipulation is constrained, managers would be more likely to inflate core earnings through classification shifting. We predict this relation as our second hypothesis:

---

[8] We find similar results for our sample of firms. The average magnitude of income-decreasing special items (scaled by sales) in the fourth quarter is 5.36 percent, compared to only 1.64 percent in the first three quarters. We also find that income-decreasing special items are reported for 31.7 percent of the fourth quarter observations, compared to only 15.6 percent of the observations for the first three quarters.

[9] The intuition behind their use of net operating assets is that the balance sheet accumulates the effects of previous accounting choices. Therefore, firms with higher than expected net operating assets are more likely to have reported optimistic earnings in the past. To validate their measure, Barton and Simko (2002) report that firms with larger net operating assets at the beginning of the current year reported more positive abnormal accruals in the previous 20 quarters. Using Dechow et al.'s (2008) Fraud Score, Omer et al. (2008) also find that the level of net operating assets provides a measure of prior earnings management.

H2: Managers shift core expenses to special items more when the ability to manipulate accruals is constrained.

Finally, we examine the extent to which managers use classification shifting as a way to meet three earnings benchmarks: analysts' forecasts, earnings of the same quarter in the prior year, and profit. In general, prior studies find evidence consistent with earnings management at each threshold (e.g., Burgstahler and Dichev 1997a; Degeorge et al. 1999; Payne and Thomas 2003; Brown and Caylor 2005). There are clear market incentives for meeting versus missing the analyst forecast threshold (e.g., Kasznik and McNichols 2002). Dechow and Skinner (2000) note that earnings management is greater when such actions allow managers to meet or beat analysts' forecasts than when they otherwise would not. On a quarterly basis, Bartov et al. (2002) find that firms meeting or beating analysts' earnings expectations enjoy a higher return over the quarter than firms that fail to meet these expectations. They find that the equity premium to firms meeting or beating expectations is only marginally affected by whether reported quarterly earnings are genuine or are the result of earnings or expectations management. Therefore, the *ex post* equity premium following quarterly earnings announcements provides an *ex ante* incentive for managers to manipulate earnings numbers.

Other studies document that analysts and investors pay more attention to "street" or pro forma earnings as defined by managers and view them as more value-relevant (Bradshaw and Sloan 2002; Bhattacharya et al. 2004; Gu and Chen 2004). Bradshaw and Sloan (2002) show that the market responds to street earnings more than traditional GAAP earnings and that managers have taken a proactive role in defining street earnings when communicating to investors. Analysts also tend to exclude special items from their forecasts. Thus, it is likely that managers take advantage of the market's focus on core earnings instead of bottom-line GAAP earnings to shift some core expense items in the income statement.

For similar reasons, managers may use classification shifting to report quarterly earnings that meet or just beat earnings of the same quarter in the prior year or that just meet or beat zero earnings. This could occur because managers have incentives (e.g., stock premium, bonuses, reputation, job security) to avoid quarterly earnings decreases and losses. In summary, we expect that managers shift core expenses to income-decreasing special items more when quarterly earnings of their firms just meet or beat (1) the analysts' consensus forecasts, (2) earnings of the same quarter in the prior year, or (3) zero earnings. Our hypothesis related to classification shifting around earnings benchmarks follows:

H3: Managers shift core expenses to special items more when quarterly earnings just meet or beat earnings benchmarks.

Degeorge et al. (1999) suggest that managers first attempt to report a profit, then to report earnings increases and, last, they are concerned with meeting or beating analysts' forecasts. Payne and Thomas (2003) find evidence consistent with earnings management mostly for the zero earnings threshold followed by the analyst forecast threshold. In contrast, Graham et al. (2005) report that managers most prefer to avoid reporting an earnings decrease. Brown and Caylor (2005) show that achieving the analyst forecast threshold has become the most important over time. Because prior research provides mixed conclusions, we have no *a priori* expectations on which of the three thresholds should show the most evidence of classification shifting.

IV. DATA, SAMPLE SELECTION, AND DESCRIPTIVE STATISTICS

We collect data for the years 1988 to 2007 from Compustat Industrial Quarterly File. Analysts' forecast data are obtained from the I/B/E/S Detail File. Each firm-quarter observation is required to have sufficient data to calculate variables in model (2) and model (3), as discussed below. Following McVay (2006), we eliminate firm-quarter observations that have annual sales of less than $1 million. This allows our sample to be as comparable to McVay's (2006) as possible. In addition, eliminating these firms reduces the extreme impact that small firms can have on the results. We also exclude firms that had a change in fiscal year, to ensure that quarterly data are comparable across years. Further, requiring a minimum of 15 observations per industry-year-quarter ensures a sufficiently large sample to estimate expected core earnings. Industry classifications are based on Fama and French (1997). The full sample consists of 132,393 firm-quarter observations. The subsample with available analysts' forecasts has 67,980 firm-quarters.

Table 1 lists the definitions of variables used in the analyses, and Table 2 provides descriptive statistics for these variables. The mean (median) of core earnings for all firm-quarters, scaled by sales, is 0.038 (0.085). The mean (median) of special items as a percentage of sales is 2.59 (0.0) percent. Income-increasing special items are set to zero and are not included in the analyses. The mean (median) of unexpected core earnings deflated by sales is 0.000 (0.005). The calculation of unexpected core earnings is discussed in detail in the following section. Table 3 presents correlations among the main variables.

TABLE 1

Variable Definitions with Corresponding Compustat Data Item Numbers

Variable		Definition
CE_q	=	Core Earnings, calculated as sales (#2) minus cost of goods sold (#30) minus selling, general, and administrative expenses (#1) in quarter q.
UE_CE_q	=	Unexpected Core Earnings, calculated as the reported core earnings (CE_q) minus the expected core earnings estimated by industry-year-quarter, excluding firm i: $$CE_q = \beta_0 + \beta_1 CE_{q-4} + \beta_2 CE_{q-1} + \beta_3 ATO_q + \beta_4 ACCRUALS_{q-4} + \beta_5 ACCRUALS_{q-1} + \beta_6 \Delta SALES_q + \beta_7 NEG\_\Delta SALES_q + \beta_8 RETURNS_{q-1} + \beta_9 RETURNS_q + \varepsilon_q.$$
$\%SI_q$	=	Special Items (#32) as a percentage of sales (#2). Income-decreasing special items are multiplied by -1, where special items are income-decreasing, and are set to 0 where special items are income-increasing.
$ACCRUALS_q$	=	Accruals, calculated as net income before extraordinary items (#8) minus cash from operations (#108).
$RETURNS_q$	=	Three-month market-adjusted return corresponding to the fiscal quarter.
ATO_q	=	Asset Turnover Ratio, defined as $SALES_q / ((NOA_q + NOA_{q-1}) / 2)$, where NOA_q, or net operating assets, is operating assets minus operating liabilities. Operating assets are calculated as total assets (#44) less cash and short-term investments (#36). Operating liabilities are calculated as total assets (#44) less total debt (#45 and #51), less book value of common and preferred equity (#55 and #56), less minority interest (#53). Average NOA is required to be positive.
$\Delta SALES_q$	=	Percentage Change in Sales, calculated as $(SALES_q - SALES_{q-4}) / SALES_{q-4}$.
$NEG\_\Delta SALES_q$	=	$\Delta SALES_q$ if the percentage change in sales is less than 0, and 0 otherwise.

TABLE 2

Descriptive Statistics

Variable	Mean	Median	Standard Deviation	25%	75%
CE_q	0.038	0.085	0.346	0.023	0.154
UE_CE_q	0.000	0.005	0.189	−0.041	0.056
$\%SI_q$	2.59%	0.00%	0.112	0.00%	0.00%
CE_{q-4}	0.025	0.086	0.391	0.023	0.154
CE_{q-1}	0.033	0.084	0.351	0.021	0.155
ATO_q	2.873	1.010	6.885	0.459	2.259
$ACCRUALS_{q-4}$	−0.093	−0.066	0.541	−0.241	0.063
$ACCRUALS_{q-1}$	−0.111	−0.067	0.519	−0.248	0.056
$\Delta SALES_q$	19.26%	8.54%	0.539	−4.39%	26.56%
$RETURNS_{q-1}$	0.002	−0.034	0.331	−0.183	0.126
$RETURNS_q$	0.001	−0.033	0.332	−0.184	0.126

See Table 1 for variable definitions. The full sample consists of 132,393 firm-quarter observations. The sample with analysts' forecast data available has 67,980 firm-quarter observations. All variables are winsorized at 1st and 99th percentile.

V. RESEARCH DESIGN AND RESULTS

Measuring Unexpected Core Earnings

To measure expected core earnings for firm i, we estimate model (2) within each industry-year-quarter excluding firm i.[10]

$$CE_q = \beta_0 + \beta_1 CE_{q-4} + \beta_2 CE_{q-1} + \beta_3 ATO_q + \beta_4 ACCRUALS_{q-4} + \beta_5 ACCRUALS_{q-1} + \beta_6 \Delta SALES_q + \beta_7 NEG\_\Delta SALES_q + \beta_8 RETURNS_{q-1} + \beta_9 RETURNS_q + \varepsilon_q. \quad (2)$$

Model (2) is a variant of the model used by McVay (2006). The intent of the core earnings expectations model is to capture the natural relation between core earnings (CE) and firm performance, leaving the unexplained portion (i.e., unexpected core earnings) as a measure of abnormal performance.[11] Lagged core earnings are included because core earnings are persistent.[12] We include core earnings from the last quarter and four quarters ago. The last quarter is closer to the current quarter, so it better captures firms' current economic environments. Thus, it may provide a better indication of normal core earnings for the current quarter. We also include core earnings four quarters ago because quarterly earnings exhibit a seasonal pattern for many firms, such that earnings of the same quarter one year ago may provide a better control for current performance.

[10] Variables are defined in Table 1.

[11] Reported core earnings are calculated as sales minus cost of goods sold and selling, general, and administrative expenses. As a sensitivity test, we also consider using I/B/E/S actual quarterly earnings in place of Compustat core quarterly earnings. While the use of I/B/E/S data substantially reduces our sample, we continue to find support for our hypotheses. So that our results are directly comparable to those in McVay (2006), we employ Compustat core earnings as our primary analysis.

[12] In Table 3, the Spearman (Pearson) correlation between CE_q and CE_{q-1} is 0.7902 (0.7351). The Spearman (Pearson) correlation between CE_q and CE_{q-4} is 0.7357 (0.6530).

TABLE 3

Spearman/Pearson Correlation Matrix

	CE_q	CE_{q-1}	CE_{q-4}	$ACCRUALS_{q-1}$	$ACCRUALS_{q-4}$	ATO_q
CE_q	1.0000	**0.7351**	**0.6530**	**−0.2750**	**−0.2806**	**0.0308**
CE_{q-1}	**0.7902**	1.0000	**0.6446**	**−0.2084**	**−0.3043**	**0.0295**
CE_{q-4}	**0.7357**	**0.6999**	1.0000	**−0.3019**	**−0.2415**	**0.0135**
$ACCRUALS_{q-1}$	**−0.2947**	**−0.2654**	**−0.3129**	1.0000	**0.3171**	**0.0381**
$ACCRUALS_{q-4}$	**−0.3055**	**−0.3441**	**−0.2722**	**0.0342**	1.0000	**0.0238**
ATO_q	**−0.0554**	**−0.0558**	**−0.0921**	**0.1460**	**0.1039**	1.0000
$RETURNS_q$	**0.1748**	**0.1503**	**0.0675**	**−0.0396**	**−0.0361**	**0.0789**
$RETURNS_{q-1}$	**0.1627**	**0.1678**	**0.0764**	**−0.0166**	**−0.0538**	**0.0819**
$\Delta SALES_q$	**0.2474**	**0.1928**	**0.0393**	**0.0800**	**−0.0094**	**0.1573**
$NEG\_\Delta SALES_q$	**−0.2621**	**−0.2150**	**−0.0900**	**−0.0486**	**0.0244**	**−0.1836**
UE_CE_q	**0.2392**	−0.0089	**0.0151**	**−0.0121**	**−0.0046**	**−0.0346**
$\%SI_q$	**−0.0616**	**−0.0327**	**−0.0163**	**−0.0339**	**−0.0685**	**−0.1000**

	$RETURNS_q$	$RETURNS_{q-1}$	$\Delta SALES_q$	$NEG\_\Delta SALES_q$	UE_CE_q	$\%SI_q$
CE_q	**0.0916**	**0.0708**	**0.0847**	**−0.2010**	**0.4476**	**−0.2293**
CE_{q-1}	**0.0678**	**0.0846**	0.0011	**−0.1468**	**0.0227**	**−0.1580**
CE_{q-4}	**0.0264**	**0.0274**	**−0.2388**	**−0.0139**	**0.0099**	**−0.1493**
$ACCRUALS_{q-1}$	**−0.0073**	0.0033	**0.0650**	**−0.0411**	−0.0020	**0.0156**
$ACCRUALS_{q-4}$	**−0.0104**	**−0.0232**	**0.0229**	**0.0132**	−0.0068	**0.0112**
ATO_q	**0.0242**	**0.0284**	**0.0488**	**−0.0764**	**−0.0077**	**−0.0434**
$RETURNS_q$	1.0000	**−0.0138**	**0.0688**	**−0.0915**	−0.0067	**−0.0755**
$RETURNS_{q-1}$	**0.0147**	1.0000	**0.0896**	**−0.1000**	−0.0052	**−0.0601**
$\Delta SALES_q$	**0.1262**	**0.1421**	1.0000	**−0.4780**	**−0.0130**	**−0.0474**
$NEG\_\Delta SALES_q$	**−0.1277**	**−0.1356**	**−0.8093**	1.0000	0.0014	**0.1107**
UE_CE_q	**−0.0277**	**−0.0115**	**−0.0972**	**0.0936**	1.0000	**−0.0942**
$\%SI_q$	**−0.0740**	**−0.0559**	**−0.0961**	**0.0951**	**−0.0119**	1.0000

Variable definitions are in Table 1. There are 132,393 firm-quarter observations. Spearman (Pearson) correlations are below (above) the diagonal. All variables are winsorized at 1st and 99th percentile. Amounts in bold are significant at the 0.01 level.

Since asset turnover (*ATO*) tends to be inversely related to profit margins (Nissim and Penman 2001) and the definition of core earnings closely parallels profit margins, we include *ATO* in the model. Sloan (1996) finds that accrual levels are an explanatory variable for future performance, so we use lagged operating accruals ($ACCRUALS_{q-1}$ and $ACCRUALS_{q-4}$) as a control for current operating performance. For reasons discussed earlier, we do not include current-period accruals. Also similar to McVay (2006), we include percentage changes in sales ($\Delta SALES$) and allow different slopes for sales increases and decreases (*NEG*). The reason for this inclusion is that, even though core earnings are scaled by sales, as sales grow, fixed costs become smaller per sales dollar. Those costs increase more when activity rises than they decrease when activity falls by an equivalent amount (Anderson et al. 2003).

In addition to the variables proposed by McVay (2006), we also add the current quarter's returns (*RETURNS*) and the previous quarter's returns to the model. The use of current quarter's

returns controls for current performance. The reason for including prior-period returns is that the market may detect deteriorating performance and decrease its expectations of core earnings prior to it being reported in the current quarter.

We obtain coefficients for model (2) using all observations in a particular industry-year-quarter excluding firm *i*. Expected core earnings for firm *i* are measured using these coefficients multiplied by the actual values of the variables in the model for firm *i*. This model controls for the macroeconomic and industry shocks as well as the seasonal effect on expected core earnings. We then obtain unexpected core earnings, calculated as the difference between reported and expected core earnings:[13]

$$UE\_CE_q = CE_q - E(CE_q). \tag{3}$$

We then estimate the relation between *UE_CE* and *%SI*:

$$UE\_CE_q = \alpha_0 + \alpha_1 \%SI_q + \varepsilon_q, \tag{4}$$

where $\%SI_q$ is special items as a percentage of sales in the current quarter. Income-decreasing special items are special items multiplied by -1, where special items are income-decreasing. Income-increasing special items are set to zero (McVay 2006).

If the estimated coefficient on $\%SI_q$ (α_1) is positive, then classification shifting has a dominating effect (i.e., as income-decreasing special items increase, unexpected core earnings increase). In other words, managers shift core expenses to special items such that reported core earnings would be above expectations. If the estimated coefficient is negative, then firm performance would be the dominating effect. In other words, when income-decreasing special items increase, unexpected core earnings actually decrease. This would be consistent with prior research findings that firms incurring large write-offs or corporate restructuring charges tend to be poor performers (Elliott and Shaw 1988; DeAngelo et al. 1994; Carter 2000).

To demonstrate how controlling for performance in model (2) affects the relation between *UE_CE* and *%SI*, we start with a model that provides no control for performance and then progressively add more controls. First, we set UE_CE_q equal to CE_q. This is the same as setting *E*(*CE*) equal to 0 and providing no control for performance. As shown in the first column of results in Table 4, the coefficient of *%SI* is -0.7099 ($t = -85.71$; $p < 0.01$). Recall that classification shifting suggests a *positive* relation between *UE_CE* and *%SI*, whereas a performance effect would predict that income-decreasing special items are a sign of poor performance (i.e., a *negative* relation between *UE_CE* and *%SI*). Thus, as expected we find strong evidence of a performance-driven relation between core earnings and income-decreasing special items.

In the second column of Table 4, we use a core earnings expectations model that relies on the assumption that expected core earnings in the current quarter are a function of only core earnings in the same quarter one year ago.

$$CE_q = \beta_0 + \beta_1 CE_{q-4} + \varepsilon_q. \tag{5}$$

Model (5) is estimated within each industry-year-quarter excluding firm *i*. Coefficient estimates, along with lagged core earnings, are used to calculate unexpected core earnings for firm *i*. We then estimate the relation between this measure of unexpected core earnings and *%SI*. The coefficient on *%SI* decreases in magnitude to -0.3531 ($t = -60.28$; $p < 0.01$).

[13] Consistent with McVay (2006), we identify $E(CE_q)$ as "expected" and UE_CE_q as "unexpected" core earnings. Since $E(CE_q)$ is partially determined by contemporaneous independent variables, these two variables are not, in fact, *ex ante* in nature and so constitute *ex post* measures of normal and abnormal core earnings.

TABLE 4

Regression of Unexpected Core Earnings on Special Items as a Percentage of Sales Using Alternative Measures of Expected Core Earnings

	Dependent Variable = $UE\_CE_q = CE_q - E(CE_q)$			
	$E(CE_q) = 0$	$E(CE_q) = f(CE_{q-4})$	$E(CE_q) = f$(McVay model)	$E(CE_q) = f$(full model)
Intercept	0.0564	0.0091	0.0063	0.0041
	(59.25)	(13.57)	(10.82)	(6.67)
$\%SI_q$	−0.7099	−0.3531	−0.2630	−0.1586
	(−85.71)	(−60.28)	(−48.06)	(−16.38)
R^2	5.26%	2.67%	1.71%	0.89%

Variable definitions are in Table 1. The full sample consists of 132,393 firm-quarter observations.

In column 1, unexpected core earnings are defined as core earnings in quarter q.

In column 2, unexpected core earnings are defined as core earnings minus expected core earnings, where expected core earnings are calculated using the coefficients from the model shown below, estimated for each industry-year-quarter, excluding firm i:

$$CE_q = \beta_0 + \beta_1 CE_{q-4} + \varepsilon_q.$$

In column 3, unexpected core earnings are defined as core earnings minus expected core earnings, where expected core earnings are calculated using the coefficients from the model shown below, estimated for each industry-year-quarter, excluding firm i:

$$CE_q = \beta_0 + \beta_1 CE_{q-4} + \beta_2 ATO_q + \beta_3 ACCRUALS_{q-4} + \beta_4 \Delta SALES_q + \beta_5 NEG\_\Delta SALES_q + \varepsilon_q.$$

In column 4, unexpected core earnings are defined as core earnings minus expected core earnings, where expected core earnings are calculated using the coefficients from the model shown below, estimated for each industry-year-quarter, excluding firm i:

$$CE_q = \beta_0 + \beta_1 CE_{q-4} + \beta_2 CE_{q-1} + \beta_3 ATO_q + \beta_4 ACCRUALS_{q-4} + \beta_5 ACCRUALS_{q-1} + \beta_6 \Delta SALES_q + \beta_7 NEG\_\Delta SALES_q + \beta_8 RETURNS_{q-1} + \beta_9 RETURNS_q + \varepsilon_q.$$

All variables are winsorized at the 1st and 99th percentile. Amounts reported are regression coefficients (with t-statistics in parentheses).

Next, we use a core earnings expectations model that approximates the one used in McVay (2006), except that it excludes current-period accruals and is estimated using quarterly data.

$$CE_q = \beta_0 + \beta_1 CE_{q-4} + \beta_2 ATO_q + \beta_3 ACCRUALS_{q-4} + \beta_4 \Delta SALES_q + \beta_5 NEG\_\Delta SALES_q + \varepsilon_q. \quad (6)$$

As shown in the third column of Table 4, the coefficient on $\%SI$ decreases in magnitude to −0.263 (t = −48.06; p < 0.01).

In the final column, we report results after using the full core earnings expectations model as specified in model (2). The estimated coefficient on $\%SI_q$ is −0.1586 (t = −16.38; p < 0.01), indicating a strong performance-dominated relation. As the results in Table 4 clearly demonstrate,

current performance and special items have a strong negative relation. However, as one provides additional controls for performance, the performance effect diminishes.[14,15]

Tests of Hypotheses

Test of H1

The first hypothesis predicts that shifting is more pervasive in the fourth quarter. As discussed in Section II, prior research provides evidence that accrual manipulation is more constrained in the fourth quarter, leaving managers to resort to alternative earnings management techniques such as income classification shifting. Thus, we expect that the relation between unexpected core earnings and special items will be more positive (or less negative) for fourth quarter observations, compared to observations for the first three quarters.

Table 5 presents the results using subsamples with observations from the fourth quarter and those from the first three quarters. The results show a less negative coefficient for the fourth quarter subsample compared to the subsample for the first three quarters.[16] Recall that a more positive (or less negative) relation between unexpected core earnings and special items is evidence of more classification shifting. Comparing the coefficients of -0.1440 for the fourth quarter and -0.1819 for the first three quarters indicates that classification shifting is more pervasive in the fourth quarter. The difference between the two estimated coefficients (0.0379) is statistically significant ($t = 4.05$; $p < 0.01$). This result provides support for H1 that classification shifting is more pervasive in the fourth quarter, even though the performance-driven effect is still the dominating factor.[17]

Test of H2

Our next hypothesis (H2) suggests that those managers who are more constrained in their ability to manipulate accruals because of previous accrual management activity are more likely to engage in classification shifting in the current quarter. To the extent that one earnings management option is not available, managers are likely to implement the next. Following Barton and Simko (2002), we measure managers' accrual manipulation constraint using net operating assets at the beginning of the quarter (scaled by sales). They demonstrate that firms that have optimistically reported accruals in prior periods tend to have higher net operating assets. We classify firms in the high (low) group when their net operating assets are above or equal to (below) the median for the

---

[14] To find direct evidence of classification shifting, ideally one would find a *positive* relation between *UE_CE* and *%SI*. Since we do not find this for the full version of the model but our results and anecdotal evidence suggests that it is occurring, this provides a fruitful avenue for future research. As the literature advances and expectations models better control for the performance effect, the relation between *UE_CE* and *%SI* is expected to become positive.

[15] To determine whether the relation between our estimate of unexpected core earnings from Model (2) and income-decreasing special items is mechanically driven by *contemporaneous* accrual special items, we estimate accrual special items as other funds from operations (#81) plus gains and losses on the sale of property, plant, and equipment and investments (#102). We split the sample into those with estimated accrual special items, scaled by sales, greater (less) than 0.5 percent (-0.5 percent). We also create a third sample with estimated accrual special items between 0.5 percent and -0.5 percent. We find the relations between *UE_CE* and *%SI* across these three groups do not differ significantly. For estimated accrual special items less than -0.5 percent (greater than 0.5 percent) [between -0.5 percent and 0.5 percent], the coefficient on *%SI* (i.e., α_1 in Model (4)) is -0.1542 (-0.1506) [-0.1736].

[16] In Table 5 and all remaining tables, differences in coefficients between Q4 and Q1–3 are tested using an indicator variable for fourth quarter observations. For example, for Table 5 we estimate the following model: $UE\_CE_q = \alpha_0 + \alpha_1 \%SI_q + \alpha_2 Q4 + \alpha_3 \%SI_q * Q4 + \varepsilon_q$, where the prediction is $\alpha_3 > 0$. In the tables, we present results separately for Q4 and Q1–3 to make the results for each subsample more evident.

[17] The increased evidence of classification shifting in the fourth quarter might be a result of increased opportunity (i.e., firms report more special items in the fourth quarter) or a result of increased incentives. However, either explanation supports H1 that classification shifting is more likely in the fourth quarter.

TABLE 5

Regression of Unexpected Core Earnings on Special Items as a Percent of Sales: Fourth Quarter (Q4) versus First Three Quarters (Q1–3)

Independent Variable	Q4	Q1–3	Test (H1): Q4 > Q1–3
Intercept	0.0070	0.0032	
	(5.41)**	(5.72)**	
$\%SI_q$	−0.1440	−0.1819	0.0379
	(−19.12)**	(−28.54)**	(4.05)**
R^2	1.08%	0.82%	
Number of Observations	33,454	98,939	

*, ** Indicates significant at the 0.05 and 0.01 levels, respectively, using a two-tailed test.

Variable definitions are in Table 1. All variables are winsorized at the 1st and 99th percentile. Amounts reported are regression coefficients (with t-statistics in parentheses).

industry-year-quarter. If classification shifting occurs, then we expect to find more evidence of it for firms with higher net operating assets (relative to the industry-quarter median) because these firms are more constrained in their ability to manipulate accruals.[18]

Results are reported in Table 6. We test H2 for the full sample and for subsamples based on Q4 and Q1–3. The subsample tests also allow us to re-examine H1 within the context of H2. The interactive effect provides a stronger test of classification shifting. If classification shifting is occurring, then we expect it to be more pervasive in fourth-quarter observations with high NOA. The results for the full sample and the subsamples support H2. For the full sample, we find that the relation between unexpected core earnings and special items is significantly less negative for firms with higher net operating assets. The difference is significant ($t = 9.88$; $p < 0.01$). The same is also true for observations in Q4 ($t = 5.52$; $p < 0.01$) and in Q1–3 ($t = 6.77$; $p < 0.01$). Furthermore, fourth quarter observations with high NOA show the most evidence in favor of classification shifting relative to the other observations. Supporting H2, we conclude that when managers' ability to manipulate accruals is constrained, classification shifting is more likely.[19]

Test of H3

For the next set of tests, we split the sample based on reported earnings relative to earnings benchmarks. Hypothesis 3 predicts that classification shifting is more prevalent in subsamples that just meet or beat the analyst forecast, one-year-ago same-quarter earnings, and zero earnings, respectively. In particular, the coefficients on $\%SI_q$ are expected to be more positive (or less negative) for subsamples with observations that have these characteristics.

[18] Of course, our use of this measure requires the joint assumption that net operating assets provide a reliable measure of prior upward accrual manipulation and that prior upward accrual manipulation reduces the ability of managers to manipulate accruals in the current period. To the extent that this joint assumption is false, inferences from our results are limited. Other studies employing net operating assets find evidence consistent with it being a measure of managers' accrual manipulation constraint (e.g., Mosebach and Simko 2005; Ettredge et al. 2007; Omer et al. 2008).

[19] While not the intended purpose of our study, these results are also consistent with a portfolio approach to managing earnings (e.g., Jiambalvo 1996; Fields et al. 2001; Lin et al. 2006; Cohen and Zarowin 2008; Cohen et al. 2008).

TABLE 6

Regression of Unexpected Core Earnings on Special Items as a Percent of Sales: High Net Operating Assets (*HighNOA*) versus Low NOA (*LowNOA*)

Independent Variable	Full Sample	Q4	Q1–3	Test (H1): Q4 > Q1–3
HighNOA	0.0130	0.0218	0.0103	
	(17.41)**	(12.05)**	(12.97)**	
LowNOA	−0.0051	−0.0084	−0.0040	
	(−6.80)**	(−4.62)**	(−5.04)**	
$\%SI_q$ * *HighNOA* (A)	−0.1153	−0.1019	−0.1441	0.0422
	(−17.83)**	(−9.44)**	(−16.55)**	(3.22)**
$\%SI_q$ * *LowNOA* (B)	−0.2063	−0.1848	−0.2305	0.0457
	(−31.49)**	(−17.70)**	(−24.67)**	(3.42)**
Test (H2): A > B	0.0910	0.0829	0.0864	
	(9.88)**	(5.52)**	(6.77)**	
R^2	1.25%	1.77%	1.07%	
Number of Observations	132,393	33,454	98,939	

*, ** Indicates significant at the 0.05 and 0.01 levels, respectively, using a two-tailed test.

Variable definitions are in Table 1. *HighNOA* (*LowNOA*) is an indicator variable that equals 1 if a firm-quarter observation has net operating assets at the beginning of the quarter above or equal to (below) the median for the industry-year-quarter, and 0 otherwise. All variables are winsorized at 1st and 99th percentile. Amounts reported are regression coefficients (with t-statistics in parentheses).

We use the I/B/E/S Detail File to determine whether firms just meet or beat analysts' forecasts. The analyst forecast is defined as the median of the last three individual analysts' forecasts after the previous quarter's earnings announcement and before the current quarter's earnings announcement.[20] Firms reporting an earnings forecast error equal to $0.00 or $0.01 are considered as just meeting or beating analysts' forecasts. These observations include approximately 26 percent of our sample. For the prior period earnings benchmark, we define firms as just meeting or beating if they report a change in reported operating income per share (Compustat data item #177) from $0.00 to $0.03, constituting approximately 17 percent of our quarterly observations. For the zero earnings benchmark, firms that report operating income per share from $0.00 to $0.04 are defined as those just meeting the threshold, which equals approximately nine percent of our quarterly observations.

Results for testing H3 are presented in Tables 7–9. In Table 7 for the full sample, the coefficient of *%SI* is −0.0956 for firms that just meet or beat the analyst forecast threshold (*JustMBF*) versus −0.1288 for other firms (*NotJustMBF*). The difference in coefficients of 0.0332 is statistically significant (t = 2.17; p = 0.03). The difference increases to 0.0581 (t = 2.01; p = 0.04) for observations in the fourth quarter and falls to 0.0251 (t = 1.36; p = 0.17) in the interim quarters.

[20] We use the I/B/E/S data unadjusted for stock splits for reasons discussed in Payne and Thomas (2003).

TABLE 7

Regression of Unexpected Core Earnings on Special Items as a Percent of Sales: Just Meet or Beat Forecasts (*JustMBF*) versus Not Just Meet or Beat Forecasts (*NotJustMBF*)

Independent Variable	Full Sample	Q4	Q1–3	Test (H1): Q4 > Q1–3
JustMBF	0.0083	0.0133	0.0069	
	(7.68)**	(4.98)**	(5.98)**	
NotJustMBF	0.0061	0.0118	0.0042	
	(9.33)**	(7.75)**	(5.93)**	
$\%SI_q$ * *JustMBF* (A)	−0.0956	−0.0533	−0.1270	0.0737
	(−6.79)**	(−1.98)*	(−7.62)**	(2.54)**
$\%SI_q$ * *NotJustMBF* (B)	−0.1288	−0.1114	−0.1521	0.0407
	(−21.57)**	(−10.78)**	(−19.84)**	(3.38)**
Test (H2): A > B	0.0332	0.0581	0.0251	
	(2.17)*	(2.01)*	(1.36)	
R^2	0.84%	0.92%	0.94%	
Number of Observations	67,980	17,563	50,417	

*, ** Indicates significant at the 0.05 and 0.01 levels, respectively, using a two-tailed test.

Variable definitions are in Table 1. *JustMBF* (*NotJustMBF*) is an indicator variable that equals 1 (0) if a firm-quarter observation has an analyst forecast error of $0.00 or $0.01, and 0 (1) otherwise. Analyst forecast error is defined as actual earnings per share as reported by I/B/E/S less the median of the last three individual analyst forecast after the previous quarter's earnings announcement and before the current quarter's earnings announcement. All variables are winsorized at 1st and 99th percentile. Amounts reported are regression coefficients (with t-statistics in parentheses).

We also note that the differences in Q4 versus Q1–3 are significant for both *JustMBF* and *NotJustMBF* firms. Overall, we conclude that finding the most evidence of classification shifting in the fourth quarter for firms that just meet or beat the analyst forecast supports H3.[21]

We then split the full sample into the subsample with firm-quarter earnings that just meet or beat prior period earnings (*JustMBP*) and the subsample with firm-quarter earnings that do not just meet or beat prior period earnings (*NotJustMBP*). Managers of firms with earnings that *ex post* just meet or beat earnings in the same quarter of previous year are expected to have had stronger incentives *ex ante* to shift core expenses to special items, especially in the fourth quarter. The results in Table 8 show that for the full sample and for each of the subsamples, the relation between unexpected core earnings and special items is less negative for *JustMBP* firms. These results are consistent with H3. We also find stronger evidence of classification shifting to just meeting or beating prior period earnings in the fourth quarter.

[21] As an additional test using analyst data, we consider the extent to which analysts exclude income-decreasing special items. We do this by comparing I/B/E/S actual earnings to reported earnings per share excluding extraordinary items from Compustat. To the extent that analysts do not incorporate income-decreasing special items in their forecasts, we expect I/B/E/S actual earnings to be greater than reported earnings. For our sample of observations with I/B/E/S data (n = 67,980), we find that approximately 23 percent have I/B/E/S earnings greater than reported earnings per share excluding extraordinary items. Furthermore, if classification shifting is meant to mislead financial statement users, then we expect to find more evidence of classification shifting when I/B/E/S earnings exclude expenses. This is what we find. When I/B/E/S earnings are greater than reported earnings, we find that the relation between unexpected core earnings and special items is less negative for the full sample ($t = 4.12$; $p < 0.01$), in Q4 ($t = 2.33$; $p = 0.02$), and in Q1–3 ($t = 3.63$; $p < 0.01$).

TABLE 8

Regression of Unexpected Core Earnings on Special Items as a Percent of Sales: Just Meet or Beat Prior Earnings (*JustMBP*) versus Not Just Meet or Beat Prior Earnings (*NotJustMBP*)

Independent Variable	Full Sample	Q4	Q1–3	Test (H1): Q4 > Q1–3
JustMBP	0.0120	0.0156	0.0113	
	(9.15)**	(4.21)**	(8.41)**	
NotJustMBP	0.0027	0.0057	0.0017	
	(4.52)**	(4.04)**	(2.72)**	
$\%SI_q$ * *JustMBP* (A)	−0.0995	−0.0741	−0.1373	0.0632
	(−6.70)**	(−3.06)**	(−6.69)**	(2.10)*
$\%SI_q$ * *NotJustMBP* (B)	−0.1630	−0.1484	−0.1860	0.0376
	(−32.74)**	(−18.23)**	(−27.09)**	(3.73)**
Test (H3): A > B	0.0635	0.0743	0.0487	
	(4.05)**	(2.91)**	(3.36)**	
R^2	0.93%	1.11%	0.88%	
Number of Observations	126,932	31,923	95,009	

*, ** Indicates significant at the 0.05 and 0.01 levels, respectively, using a two-tailed test.

Variable definitions are in Table 1. *JustMBP* (*NotJustMBP*) is an indicator variable that equals 1 (0) if a firm-quarter observation has a reported change in operating income per share (Compustat data item #177) in the current quarter compared to the same quarter in the prior year between $0.00 and $0.03, and 0 (1) otherwise. All variables are winsorized at 1st and 99th percentile. Amounts reported are regression coefficients (with t-statistics in parentheses).

Table 9 reports results for firms that just meet or beat the zero earnings threshold (*JustMBZ*) versus other firms (*NotJustMBZ*). For the full sample, the coefficient on *%SI* for *JustMBZ* firms is positive (0.0304) and significant ($t = 1.96$; $p = 0.05$), while the coefficient on *%SI* for *NotJustMBZ* firms is significantly negative (−0.1762; $t = -36.60$; $p < 0.01$). The difference is significant ($t = 12.71$; $p < 0.01$). We also find significant differences in Q4 and in Q1–3. For *JustMBZ* firms, the coefficient on *%SI* is positive in Q4 (0.0400) and negative in Q1–3 (−0.0039), but the difference is not significant ($t = 1.39$; $p = 0.16$), probably due to lower power tests resulting from the small number of observations that just meet or beat the zero earnings threshold. The difference between Q4 and Q1–3 for *NotJustMBZ* is significantly positive ($t = 3.73$; $p < 0.01$). Overall, the evidence supports classification shifting for firms that just meet or beat the zero earnings benchmark, again, supporting H3.

As a final test, we consider the three earnings thresholds jointly. That is, we test for income classification shifting when firm-quarter observations just meet or beat *any* of the three earnings thresholds. This joint test presumably contains less measurement error because evidence in favor of classification shifting is found for each of the three earnings thresholds in Tables 7–9. For example, in our test of the zero earnings threshold in Table 9, the explicit assumption is that firms that are classified as not just meeting or beating the benchmark (*NotJustMBZ*) are less likely to have shifted core expenses. However, these firms may just meet or beat the analyst forecast (prior period earnings) threshold. These firms may shift in response to these alternative earnings benchmarks, but our test in Table 9 would not properly test this.

Because we require information for all earnings thresholds jointly, the sample size reduces to 66,034 firm-quarter observations. We first note that for this smaller sample, approximately 12.6

TABLE 9

Regression of Unexpected Core Earnings on Special Items as a Percent of Sales: Just Meet or Beat Zero Earnings (*JustMBZ*) versus Not Just Meet or Beat Zero Earnings (*NotJustMBZ*)

Independent Variable	Full Sample	Q4	Q1–3	Test (H1): Q4 > Q1–3
JustMBZ	0.0388	0.0514	0.0355	
	(22.46)**	(11.41)**	(19.76)**	
NotJustMBZ	0.0005	0.0031	−0.0003	
	(0.90)	(2.30)*	(−0.47)	
$\%SI_q$ * *JustMBZ* (A)	0.0304	0.0400	−0.0039	0.0439
	(1.96)*	(1.57)	(−0.18)	(1.39)
$\%SI_q$ * *NotJustMBZ* (B)	−0.1762	−0.1622	−0.1980	0.0358
	(−36.60)**	(−20.65)**	(−29.76)**	(3.67)**
Test (H3): A > B	0.2066	0.2022	0.1941	
	(12.71)**	(7.59)**	(8.63)**	
R^2	1.45%	1.78%	1.32%	
Number of Observations	132,393	33,454	98,939	

*, ** Indicates significant at the 0.05 and 0.01 levels, respectively, using a two-tailed test.

Variable definitions are in Table 1. *JustMBZ* (*NotJustMBZ*) is an indicator variable that equals 1 (0) if a firm-quarter observation has a reported operating income per share in the current quarter between $0.00 and $0.04, and 0 (1) otherwise. All variables are winsorized at 1st and 99th percentile. Amounts reported are regression coefficients (with t-statistics in parentheses).

(26.8) [33.3] percent of the observations that are classified as *not* just meeting or beating the analyst forecast (prior period earnings) [zero earnings] threshold actually just meet or beat one of the other two thresholds. Thus, we know that our tests in Tables 7–9 contain some measurement error and inferences could be affected. The results of our joint tests are reported in Table 10. We find for the full sample, for Q4, and for Q1–3 that the coefficient on *%SI* is significantly less negative for firms that just meet or beat any earnings threshold (*JustMB*) than for other firms (*NotJustMB*). We also note that the coefficient on *%SI* is significantly less negative in Q4 than in Q1–Q3 for *JustMB* firms and for *NotJustMB* firms. The results in Table 10 provide broad support for H3.

VI. CONCLUSION

This study addresses the issue of whether and when managers use classification shifting to manage core earnings. Prior research provides evidence that investors and analysts tend to fixate on core earnings (Bradshaw and Sloan 2002; Gu and Chen 2004; Kinney and Trezevant 1997) and that earnings reported higher in the income statement tend to have higher valuation multiples (Lipe 1986). Thus, the extent to which managers shift core expenses to special items is important for understanding how managers may take advantage of the market's fixation on core earnings numbers.

McVay (2006) documents a *positive* relation between unexpected core earnings and income-decreasing special items, concluding that managers shift core expenses to special items to improve core earnings. However, McVay (2006) notes that this finding does not hold (and in fact becomes *negative*) when contemporaneous total accruals (which include accrual special items) are excluded from core earnings expectations model. Thus, there is some debate of the extent to which man-

TABLE 10

Regression of Unexpected Core Earnings on Special Items as a Percent of Sales: Just Meet or Beat Any of the Three Earnings Benchmarks (*JustMB*) versus Not Just Meet or Beat Any of the Three Earnings Benchmarks (*NotJustMB*)

Independent Variable	Full Sample	Q4	Q1–3	Test (H1): Q4 > Q1–3
JustMB	0.0108	0.0196	0.0083	
	(11.84)**	(8.56)**	(8.63)**	
NotJustMB	0.0040	0.0078	0.0027	
	(5.56)**	(4.64)**	(3.49)**	
$\%SI_q * JustMB$ (A)	−0.0812	−0.0445	−0.1140	0.0695
	(−6.96)**	(−2.02)*	(−8.19)**	(2.90)**
$\%SI_q * NotJustMB$ (B)	−0.1355	−0.1185	−0.1569	0.0384
	(−21.42)**	(−10.90)**	(−19.15)**	(3.01)**
Test (H3): A > B	0.0543	0.0740	0.0429	
	(4.09)**	(3.00)**	(2.65)**	
R^2	0.93%	1.13%	0.98%	
Number of Observations	66,034	16,995	49,039	

*, ** Indicates significant at the 0.05 and 0.01 levels, respectively, using a two-tailed test.

Variable definitions are in Table 1. *JustMB* (*NotJustMB*) is an indicator variable that equals 1 (0) if a firm-quarter observation has any of the following: an analyst forecast error of $0.00 or $0.01, a reported change in operating income per share (Compustat data item #177) in the current quarter compared to the same quarter in the prior year between $0.00 and $0.03, or a reported operating income per share in the current quarter between $0.00 and $0.04, and 0 (1) otherwise. All variables are winsorized at 1st and 99th percentile. Amounts reported are regression coefficients (with t-statistics in parentheses).

agers shift core expenses to special items. McVay (2006) calls for research to improve the core earnings expectations model and to provide additional cross-sectional tests of classification shifting.

Using quarterly data, we first extend McVay's (2006) core earnings expectations model by dropping current-period accruals and including additional controls for performance. We develop alternative measures for unexpected core earnings based on different models for expected core earnings. We then examine the relation between unexpected core earnings and income-decreasing special items (measured as a positive number). Without sufficient control for performance, we first observe a strong negative relation, as opposed to a positive relation that indicates classification shifting. However, this negative relation significantly diminishes with additional performance controls.

Next, we provide additional cross-sectional tests to evaluate increased classification shifting in specific settings. First, we show that classification shifting is more prevalent in the fourth quarter than in interim quarters. This finding is consistent with managers adopting alternative earnings management techniques in the fourth quarter (Cohen et al. 2009) and that it is more difficult to achieve earnings thresholds using accrual manipulation in the fourth quarter (Brown and Pinello 2007). Second, we document that classification shifting is more prominent when managers appear to be constrained in their ability to manipulate current-period accruals because of prior upward accrual manipulation (Barton and Simko 2002). This result is consistent with the reasoning that managers have a portfolio of choices for manipulating reported performance. To the extent that one option for manipulating earnings is limited, managers will turn to the next. While

accrual manipulation and real activities management have been extensively investigated in the literature, income classification shifting has received relatively little attention and represents a viable opportunity for managers to manipulate reported earnings.

In addition, we find greater evidence of classification shifting for samples of firms that just meet or beat analysts' forecasts, one-year-ago same-quarter earnings, and zero earnings. These results are consistent with managers attempting to avoid market penalties (Barth et al. 1999; Bartov et al. 2002; Kasznik and McNichols 2002) and other negative consequences associated with missing earnings targets. These results are also consistent with managers having more incentive to shift income when investors are likely giving more weight to earnings performance (e.g., Hayn 1995; Burgstahler and Dichev 1997b; Barth et al. 1998; Collins et al. 1999; Barua et al. 2006).

Clearly, the extent, timing, and incentives under which managers engage in classification shifting is an important area of research. To the extent that managers attempt to fool investors by shifting core expenses to special items, it is important for researchers to better understand the effects on investors, auditors, regulators, and other market participants. Because this area of research is relatively new, there is ample opportunity for future research to explore the settings under which classification shifting occurs and to improve upon models to measure expected core earnings or core expenses.[22] Although we improve upon the model in McVay (2006), we have not been able to fully control for performance. Future research may consider further advancements of the model to facilitate research on classification shifting.

REFERENCES

Anderson, M., R. Banker, and S. Janakiraman. 2003. Are selling, general, and administrative costs "sticky"? *Journal of Accounting Research* 41 (1): 47–63.

Baber, W. R., P. M. Fairfield, and J. A. Haggard. 1991. The effect of concern about reported income on discretionary spending decisions: The case of research and development. *The Accounting Review* 66: 818–829.

Barth, M., W. Beaver, and W. Landsman. 1998. Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics* 25: 1–34.

———, J. Elliott, and M. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (Autumn): 387–413.

Barton, J., and P. J. Simko. 2002. The balance sheet as an earnings management constraint. *The Accounting Review* 77 (Supplement): 1–27.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (June): 173–204.

Barua, A., J. Legoria, and J. S. Moffitt. 2006. Accruals management to achieve earnings benchmarks: A comparison of pre-managed profit and loss firms. *Journal of Business Finance & Accounting* 33 (5–6): 653–670.

Bhattacharya, N., E. L. Black, T. E. Christensen, and R. D. Mergenthaler. 2004. Empirical evidence on recent trends in pro forma reporting. *Accounting Horizons* 18 (1): 27–43.

Bradshaw, M., and R. Sloan. 2002. GAAP versus the street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (1): 41–66.

———, S. Richardson, and R. Sloan. 2006. The relation between corporate financing activities, analysts'

[22] For example, an additional setting in which managers may be motivated to engage in income classification shifting is prior to issuance of equity or debt. As a preliminary investigation into this issue, we test for increased classification shifting in the current quarter when firms issue equity or debt in the following fiscal year, using external financing measures proposed by Bradshaw et al. (2006). We find no evidence of differential income classification shifting but caution that our tests are preliminary and could be expanded in many ways.

forecasts and stock returns. *Journal of Accounting and Economics* 42 (1–2): 53–85.

Brown, L., and M. Caylor. 2005. A temporal analysis of quarterly earnings thresholds: Propensities and valuation consequences. *The Accounting Review* 80 (2): 423–440.

———, and A. Pinello. 2007. To what extent does the financial reporting process curb earnings surprise games? *Journal of Accounting Research* 45 (5): 1–35.

Burgstahler, D., and I. Dichev. 1997a. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (1): 99–126.

———, and ———. 1997b. Earnings adaptation and equity value. *The Accounting Review* 72: 187–215.

———, J. Jiambalvo, and T. Shevlin. 2002. Do stock prices fully reflect the implications of special items for future earnings? *Journal of Accounting Research* 40 (3): 585–612.

Bushee, B. 1998. The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* 73: 305–333.

Carter, M. E. 2000. Operating performance following corporate restructurings. Working paper, Columbia University.

Cohen, D., A. Dey, and T. Lys. 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods. *The Accounting Review* 83 (3): 757–787.

———, and P. Zarowin. 2008. Accrual-based and real earnings management activities around seasoned equity offerings. Working paper, New York University.

———, R. Mashruwala, and T. Zach. 2009. The use of advertising activities to meet earnings benchmarks: Evidence from monthly data. Working paper, New York University.

Collins, D., M. Pincus, and H. Xie. 1999. Equity valuation and negative earnings: The role of book value of equity. *The Accounting Review* 74: 29–62.

Cornell, B., and W. Landsman. 1989. Security price response to quarterly earnings announcements and analysts' forecasts revisions. *The Accounting Review* 64: 680–692.

DeAngelo, H., L. DeAngelo, and D. Skinner. 1994. Accounting choice in troubled companies. *Journal of Accounting and Economics* 14 (1): 51–89.

Dechow, P., and D. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (2): 235–250.

———, W. Ge, C. Larson, and R. Sloan. 2008. Predicting material accounting manipulations. Working paper, University of California, Berkeley, University of Washington, and University of Michigan.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *The Journal of Business* 72 (January): 1–33.

Elliott, J., and W. Shaw. 1988. Write-offs as accounting procedures to manage perceptions. *Journal of Accounting Research* 26 (Supplement): 91–119.

Ettredge, M., S. Scholz, K. Smith, and L. Sun. 2007. The association between misstatements and bloated balance sheets. Working paper, University of Kansas.

Fairfield, P., R. Sweeney, and T. Yohn. 1996. Accounting classification and the predictive content of earnings. *The Accounting Review* 71 (3): 337–355.

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153–193.

Fields, T., T. Lys and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31 (1–3): 255–308.

Francis, J., J. Hanna, and ———. 1996. Causes and effects of discretionary asset write-offs. *Journal of Accounting Research* 34 (Supplement): 117–134.

Graham, J. R., C. R. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40 (1–3): 3–73.

Gu, Z., and T. Chen. 2004. Analysts' treatment of nonrecurring items in street earnings. *Journal of Accounting and Economics* 38 (Conference Issue): 129–170.

Gunny, K. 2005. What are the consequences of real earnings management. Working paper, University of Colorado at Boulder.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20: 125–164.

Healy, P., and J. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (4): 365–383.

Jiambalvo, J. 1996. Discussion of: Causes and consequences of earnings manipulation: An analysis of firms

subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (1): 37–47.

Kasznik, R., and M. McNichols. 2002. Does meeting expectations matter? Evidence from analysts' forecasts revisions and share prices. *Journal of Accounting Research* 40 (3): 727–759.

Kinney, M., and R. Trezevant. 1997. The use of special items to manage earnings and perceptions. *Journal of Financial Statement Analysis* (Fall): 45–53.

Lin, S., S. Radhakrishnan, and L. Su. 2006. Earnings management and guidance for meeting or beating analysts' earnings forecasts. Working paper, California State University, Fresno, The University of Texas at Dallas, and Hong Kong Polytechnic University.

Lipe, R. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (Supplement): 37–64.

McVay, S. 2006. Earnings management using classification shifting: An examination of core earnings and special items. *The Accounting Review* 81 (May): 501–531.

Mosebach, M., and P. Simko. 2005. Earnings management and the emergence of profitability. Working paper, University of Illinois.

Nissim, D., and S. Penman. 2001. Ratio analysis and equity evaluation: From research to practice. *Review of Accounting Studies* 6 (1): 109–154.

Omer, T., M. Shelley, and A. Thompson. 2008. Balance sheet bloat, earnings management, and Staff Accounting Bulletin No. 108. Working paper, Texas A&M University.

Payne, J., and W. Thomas. 2003. The implications of using stock-split adjusted I/B/E/S data in empirical research. *The Accounting Review* 78: 1049–1067.

Roychowdhury, R. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42: 335–370.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1325–1346

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1325

Joint Effects of Principles-Based versus Rules-Based Standards and Auditor Type in Constraining Financial Managers' Aggressive Reporting

Karim Jamal
University of Alberta

Hun-Tong Tan
Nanyang Technological University

ABSTRACT: Managers sometimes implement accounting standards (such as the lease standard) opportunistically to move debt off balance sheet. Regulators and standard-setters are considering the adoption of principles-based accounting standards to reduce such opportunism. We report the results of an experiment in which experienced financial managers, with incentives to structure a transaction off balance sheet, take a reporting position on how a lease is to be reported. We manipulate the type of accounting standards (principles-based, rules-based) and the type of auditor (principles-oriented, rules-oriented, or client-oriented). Results show that for a rules-based standard, auditor-type does not influence participants' propensity to report the transaction off balance sheet. However, for a principles-based standard, auditor-type matters in that this propensity is lowest when the auditor is principles-oriented as opposed to rules- or client-oriented. Our results suggest that a move toward more principles-based standards is likely to result in improved financial reporting quality only when there is a corresponding shift in auditors' mindsets toward being more principles-oriented.

Keywords: *principles-based standards; rules-based standards; auditor-type; off balance sheet debt; transaction structuring.*

We thank Steve Glover, John Ferguson, The FEI National, Edmonton and Calgary Chapter, The Institute of Chartered Accountants of Alberta, and CMA Alberta for their assistance in obtaining participants for the study. We appreciate helpful comments from Laureen Maines (editor), Steve Kachelmeier (senior editor), two anonymous referees, as well as workshop participants at the University of Alberta, Chinese University of Hong Kong, University of Hawaii, Peking University, University of Toronto, Tsinghua University, and Yale University. We also thank Wei Chen, Krista Fiolleau, Kris Hoang, Elaine Wang, and Bo Zhou for research assistance, and the Social Science and Humanities Research Council of Canada (SSHRC) and Nanyang Technological University for their financial support of this study.

Editor's note: Accepted by Laureen Maines.

Submitted: July 2008
Accepted: December 2009
Published Online: June 2010

I. INTRODUCTION

In this study, we investigate the joint effects of principles- versus rules-based standards and the type of audit partner overseeing the audit on the proclivity of managers to undertake aggressive reporting. Specifically, we examine whether the extent to which principles-based (as opposed to rules-based) standards mitigate managers' aggressive reporting is contingent on auditor-type. This issue is important because of alleged benefits of principles-based standards, the intention of the Securities and Exchange Commission (SEC) to move toward a principles-based standards framework (SEC 2003), and concerns that auditors have become rules-oriented (Financial Accounting Standards Board [FASB] 2004; Essaides 2006).

We adopt the definitions in the SEC 2003 report, *Study Pursuant to Section 108(d) of the Sarbanes-Oxley Act of 2002 on Adoption by the United States Financial Reporting System of A Principles-Based Accounting System*, and consider standards to be principles-based if they are based on a consistently applied conceptual framework with a clearly stated accounting principle, have an appropriate but not excessive level of implementation guidance, and possess few bright-line tests as well as exceptions to the principles underlying the standard.[1] One salient feature of principles-based standards is that, as compared to rules-based standards, they narrow reporting options since reliance on an accounting principle tied to a conceptual framework likely leads to a preferred accounting treatment and reduces exceptions (SEC 2008). Indeed, commenting on the benefits of principles-based standards, regulators have indicated that "with management and auditors held accountable for fulfilling the narrowly circumscribed substantive accounting objectives built into objectives-oriented standards, we believe that the degrees of freedom to achieve 'desired' accounting results is minimized relative to either a rules-based or a principles-only approach" (SEC 2003, Section 1(c)). Whether this is the case is an empirical question. To investigate our research question, we use the context of existing lease standards as an example of rules-based standards, and a new proposed lease accounting treatment as an example of principles-based standards.

While there has been much debate about the effects of principles- versus rules-based standards, less attention has been paid to the role of the auditor in implementing generally accepted accounting principles (GAAP). We posit that auditor-type matters in the implementation of principles- versus rules-based standards. Although auditor-type varies in many ways, we focus on three—whether the auditor is relatively more rules-oriented, principles-oriented, or client-oriented. These three types of auditors likely exist regardless of GAAP (principles- versus rules-based) regimes. We consider a rules-oriented auditor to be one with a relatively greater concern toward adhering to well-specified rules rather than capturing the underlying substance of the transaction. Compliance with the letter of the rule rather than fair presentation would be the orientation of such an auditor. An excessive emphasis on following the letter of the rule is of concern to regulators and standard-setters in the U.S. (Essaides 2006; Scott 2006). A principles-oriented auditor is one who focuses on the underlying substance of a transaction rather than literally following the rules. This fits the profile of the kind of auditors who regulators and standard-setters would like to see (Millman 2005), and regulators have indicated that with principles-based standards, "there is a need to train preparers and auditors in understanding the substance of the class of transactions"

[1] The SEC 2003 report also refers to principles-based standards as objectives-oriented standards. It notes that there is current confusion between principles-based standards (which the SEC advocates) and principles-only standards, and states that two distinguishing features of principles-only standards are that they involve little implementation guidance, and require significant judgment in their implementation without providing a sufficient structure to frame these judgments. In fact, the SEC report indicates that properly constructed principles-based standards may involve less judgment than either rules-based or principles-only standards.

(SEC 2003, Section 3(i)). Finally, even in a post-Sarbanes-Oxley Act (SOX) world, it is likely that client-oriented auditors exist as motivations to please clients cannot be completely eliminated (Moore et al. 2006).

We posit that the joint consideration of both GAAP-type (principles-based, rules-based) and auditor-type (principles-, rules-, or client-oriented) is necessary to limit managers' propensity to engage in aggressive accounting reporting practices. Our underlying premise is that accounting standards serve as a source of negotiation power, in that they justify and legitimatize a preferred stand by an auditor or manager (French and Raven 1959; Fiske and Berdahl 2007). A rules-based standard vests negotiation power with the client insofar as the rules in the accounting standard permit the client's preferred stand, even if doing so does not reflect the underlying economics of the transaction. An auditor who is more principles-oriented and wishes to reflect the economic substance of a client's transaction in the financial statements is constrained from doing so because he/she has less negotiation power in this situation. Thus, with rules-based standards, the nature of the auditor does not constrain the client's propensity to engage in aggressive reporting. In contrast, with a principles-based standard, auditor-type plays a part. The client knows that given a principles-based standard, the principles-oriented auditor is more likely and better able to oppose aggressive reporting. Thus, with a principles-based standard, the client's propensity to engage in aggressive reporting is more likely to be reduced with a principles-oriented auditor than either a rules-oriented or client-oriented auditor.

We conduct an experiment to investigate how auditor-type and GAAP-type jointly influence financial officers' proclivity to engage in aggressive reporting. We consider accounting reporting to be aggressive if it is biased toward an outcome consistent with management's incentives—specifically, reporting a lease-like transaction off balance sheet rather than as a capital lease (debt) in our experiment. We employ an experimental approach because archival data either do not exist within current accounting regimes (e.g., U.S. GAAP has not converted to a principles-based regime) or entail confounds when accounting judgments and market reactions are compared between two jurisdictions.

The participants in our experiment are chief financial officers (CFOs) and highly experienced CFO-type financial managers. In our experiment, the financial manager has an objective of carrying out a lease-like transaction without reporting additional debt on the balance sheet. We examine a setting where the proposed transaction has not yet been carried out, thus giving the manager an opportunity to structure the transaction before interacting with the auditor. We manipulate two between-subjects variables: the type of GAAP (principles-based or rules-based) and the type of audit partner (principles-oriented, rules-oriented, or client-oriented) with whom the manager anticipates he/she will interact. Participants are asked to assume that they operate under either existing Canadian GAAP (that is rules-based) or a proposed International Accounting Standards Board (IASB) guidance consistent with the recommendation of the G4+1 (that is principles-based).[2] The nature of the proposed transaction is such that the bright-line criteria under the rules-based standard permit it to be reported off balance sheet (as desired by the manager), thus serving as a baseline condition in which we expect the manager to report the transaction off balance sheet as an operating lease, and the auditor to have no effect on the aggressiveness of financial reporting. In contrast, with a principles-based standard, we expect a GAAP-type × auditor-type interaction in that the likelihood of aggressive financial reporting is lower when the auditor is principles-oriented than when the auditor is either rules-oriented or client-oriented.

[2] The G4+1 was an association of the standard-setting bodies of Australia, Canada, New Zealand, U.K., and U.S. with the International Accounting Standards Committee (IASC) participating as an observer. The G4+1 was disbanded on February 1, 2001 after the formation of the International Accounting Standards Board (IASB).

Our results indicate that when GAAP is rules-based, the type of audit partner (even one who is principles-oriented) does not influence the managers' propensity to take the lease off balance sheet. However, when GAAP is principles-based, auditor-type matters—the manager's propensity to take the lease off balance sheet is lower when the audit partner is principles-oriented compared to the situation when the audit partner is either rules-oriented or client-oriented. In fact, this propensity to take the lease off balance sheet is lowest when a principles-based accounting standard is adopted *and* the auditor enforces the substance of the transaction (i.e., the auditor is principles-oriented), compared to any other combination of GAAP-type or auditor-type.

Our study makes a contribution to the literature by providing what we believe is the first evidence on how principles-based versus rules-based standards and the type of auditor jointly constrain managers' aggressive reporting. Our findings provide useful evidence on the debate relating to the efficacy of principles- versus rules-based standards. We also provide *ex ante* evidence to help standard-setters evaluate the effectiveness of a proposed lease accounting standard. While the FASB and IASB believe that this narrowing of lease-reporting options (to only capital leases) results in reduced opportunities to engage in opportunistic transaction structuring (FASB/IASB 2009), our results suggest that this desired outcome is dependent on the orientation of auditors. At a broader level, our study also informs extant research on auditor-client negotiation, which has largely focused on the reporting of transactions that have already occurred (e.g., Sanchez et al. 2007; Tan and Trotman 2010). Our results suggest a more complex picture than that assumed in the current literature—an anticipated negotiation with an auditor can influence the way that the client structures a transaction that has yet to occur.

Section II discusses the previous literature and develops our hypotheses. Section III describes the experiment. Section IV presents the results. Section V summarizes the findings and presents a discussion of implications and future research opportunities.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

Current Lease Accounting Standards

In this study, we examine a context in which management of an oil and gas company is contemplating a leasing transaction, and prefers to classify the lease as an operating lease (to keep the debt off balance sheet). Current lease accounting standards (Statement of Financial Accounting Standards (SFAS) No. 13 in the U.S., International Accounting Standard (IAS) 17 by IASB, and the Canadian Institute of Chartered Accountants Handbook (CICA HB) Section 3065 in Canada) have been considered to be examples of rules-based standards (Tweedie 2007; SEC 2008). They identify four tests to determine whether a lease is classified as operating or capital: ownership transfer, lease term for a majority of asset's useful life, present value of commitments accounts for substantial part of asset's fair value, and bargain purchase option. The primary concern about current lease accounting standards is that they permit two classes of leases—an operating lease that is off balance sheet and a capital lease that is on the balance sheet—even though in both cases, the right to use the leased item and obligation to pay rentals give rise to assets and liabilities (Nailor and Lennard 2000; FASB/IASB 2009). Hence, the existing standards are inconsistent with the underlying conceptual framework (FASB/IASB 2009). Furthermore, the four criteria used in these standards are bright-line tests that can be interpreted as conditions for classifying leases (SEC 2008, 25). These attributes render current lease standards more rules-based than principles-based.

In February 2000, the standard-setters of the G4+1 countries (with the IASC as an observer) issued a report that sought to close the loophole caused by preparers gaming the distinction between capital and operating leases in the current lease standards (Nailor and Lennard 2000). The G4+1 report recommends that organizations report in the balance sheet the fair value of all assets and liabilities contained in a lease agreement, a stand that has found support by academics

(Ryan et al. 2001) and standard-setters (Tweedie 2007). In effect, this recommendation results in all leases being classified as capital leases, and eliminates consideration of the traditional four lease-classification criteria. More recently, the IASB and FASB issued a joint discussion paper making a similar proposed stand and arguing that current lease standards are conceptually flawed and inconsistent (FASB/IASB 2009). The proposal is for a narrow implementation approach—that is, the recommendation is made in the context of existing lease standards. Hence, with respect to IASB standards, the existing guidance contained in IAS 17 applies, except that the four indicators listed in IAS 17 would no longer apply. For ease of exposition, we call this the IAS 17-plus guidance, and consider it to be principles-based since it specifies a clear principle (all leases are debt that should be reported on the balance sheet) and eliminates tests that could justify different reporting for similar transactions. In contrast, the CICA HB 3065 guidance, like existing U.S. lease standards, is rules-based in that it contains "bright-line" criteria and is inconsistent with the conceptual framework (FASB/IASB 2009).

In our study, the structure of the proposed lease is such that, based on the "bright-line" criteria listed in CICA HB 3065, the stated lease term and present value of commitments enables the transaction to be classified as an operating lease and be taken off the balance sheet, although there is a penalty after 25 years if the company does not renew the lease. This penalty creates some ambiguity that could be used to argue in favor of capital or operating lease treatment. In contrast, application of IAS 17-plus suggests that the substance of the transaction is a lease, resulting in the recognition of the assets and liabilities of the lease in the balance sheet (i.e., the transaction is considered a capital lease). However, there are several ways to get around this prescribed treatment. There are scope limitations in that extractive assets (e.g., oil and gas) as well as executory contracts are excluded, and there are transition provisions and other definitions that can be used to get around the treatment prescribed in the standard. To achieve a desired reporting treatment, the manager can also alter the terms of the lease, or take the transaction off balance sheet by making the transaction something other than a lease (e.g., by getting a utility to build the power plant and then entering into a long-term purchase contract to buy power). Thus, a manager could justify either on or off balance sheet treatment under either a rules-based or principles-based standard.

Principles-Oriented, Rules-Oriented, and Client-Oriented Auditors

We argue that discussions of the effects of principles- and rules-based standards that do not also consider the effects of auditor-type are incomplete. We identify three auditor-types (principles-oriented, rules-oriented, and client-oriented) that correspond to profiles of auditors discussed by standard-setters and regulators.

An auditor intent on enforcing an accounting standard can do so by being more strict and firm in his/her stand that the underlying principles of the standard are adhered to and that the economic substance of the transaction is captured, even if the specific rules in the standard are not met. This corresponds to the profile of an "ideal" auditor envisioned by some regulators and standard-setters (Millman 2005; SEC 2003). We consider this auditor to be principles-oriented.

Alternatively, the auditor can be relatively strict and firm in ensuring that the specific rules underlying the standard are adhered to, with less consideration for the underlying principles of the standard or the economic substance of the transaction.[3] We consider this auditor to be rules-oriented. Anecdotal evidence from discussions with senior partners from public accounting firms, as well as public positions adopted by regulators (Taub 2005; Public Company Accounting Oversight Board [PCAOB] 2007), suggests that auditors who fit the profile of being rules-oriented are

[3] These two dimensions are not necessarily opposing, and an auditor can clearly do both when the rules and principles do not conflict.

relatively common (FASB 2004, 6–7). In fact, one of the main concerns about SOX reforms and PCAOB inspections is that the auditor had become risk-averse and started to interpret GAAP literally rather than focusing on the underlying substance of the transaction (Grout et al.1994; Essaides 2006; Scott 2006; PCAOB 2007). This rules-oriented auditor-type has been blamed for multiple restatements by companies, inadequate fraud detection work (PCAOB 2007), excessive internal controls testing to achieve SOX 404 certification, and unnecessary implementation costs for SOX reforms (Millman 2005).[4]

Finally, the auditor could be client-oriented in terms of having a proclivity to take the side of the client. Prior research indicates that auditors have incentives to please clients given the auditor-client fee-dependence relationship (Moore et al. 2006), and that auditors exploit ambiguities in accounting standards to justify allowing clients to engage in opportunistic interpretations of GAAP (Hackenbrack and Nelson 1996). While regulators' concerns about auditor independence and audit quality in recent years may constrain the extent to which auditors are pro-client, questionnaire results also indicate that CFOs generally feel that they "win" when there is a dispute with the auditor (e.g., Gibbins et al. 2005). Recent empirical-archival research using post-SOX data provides evidence that audit quality continues to be influenced by incentives arising from the auditor-client fee-dependence relationship (Lim and Tan 2008).

Accounting Standards as Source of Negotiation Power

Financial statements are considered to be the outcome of negotiations between auditors and client (Antle and Nalebuff 1991; Gibbins et al. 2001), and negotiation research in psychology suggests that both the stand of the auditor with respect to an accounting issue and auditor negotiation power (relative to the manager) are major determinants of such negotiation outcomes (Bacharach and Lawler 1981; Pruitt and Carnevale 1993). An auditor who is either principles-oriented or rules-oriented can be considered to have a pre-negotiation stand with respect to the treatment of an accounting issue (whether by adherence to principles or rules). Technically, an auditor's stand determines the negotiated outcome as the auditor exercises the final prerogative in terms of what is ultimately recorded in audited financial statements. However, many issues in accounting involve a range of acceptable outcomes that are subjected to discussions or negotiations with client managers, and the extent to which an auditor's initial preferred stand prevails during auditor-client negotiation is influenced by the auditor's negotiation power relative to the manager.

Accounting standards serve as a source of this negotiation power, in that they justify and legitimatize a preferred stand by an auditor or manager (French and Raven 1959; Fiske and Berdahl 2007). Although auditors have the final say in terms of the treatment of an accounting issue, the auditor's stand has to be justifiable and is constrained by applicable accounting standards. Thus, accounting standards serve as a source of negotiation power for the manager. Finding authoritative support for one's position is rated as the most important factor affecting auditor-client negotiations (Gibbins et al. 2001). In addition, from the manager's perspective, there is a marked difference in terms of the perceived norms attributable to a principles- versus rules-based standard,

[4] The most recent example in which auditors are cited as being excessively cautious and misapplying the spirit of the rules is in the conduct of SOX 404 internal control audits. The PCAOB replaced its Auditing Standard No. 2 with Auditing Standard No. 5 on July 25, 2007 to induce auditors to use more (risk-based) judgment when conducting internal control audits. This change in auditing standard occurred after widespread complaint that the auditors had narrowly applied the guidance in Auditing Standard No. 2 and caused firms to incur excessive costs (Weirich and Rouse 2007). With respect to SOX 404 audits, the actions of the auditor were seen as being defensive and inconsistent with the spirit of the auditing standard.

and psychology research indicates that such norms influence managers' judgments (Cialdini 2007; Coleman 1990). For a principles-based standard, managers may infer that the spirit of the standard is to record the underlying economics of a transaction. For a rules-based standard, managers may perceive that, given the specificity of the rules, it is more legitimate to patently apply explicit exceptions or loopholes in the rules to achieve their preferred way of recording a transaction.

Joint Effect of Standards and Auditor-Type

The discussion above suggests that, in situations where a manager disagrees with a preferred stand by the auditor, the manager can use the power associated with accounting standards to initiate negotiations and counter the auditor's stand (e.g., Fiske and Berdahl 2007; Magee et al. 2007). Thus, although the orientation of an auditor (e.g., principles- versus rules-oriented) likely determines the stand the auditor takes with respect to an accounting issue, the nature of the accounting standard can undermine or support the degree to which the auditor can enforce his/her preferred stand.

To the extent that the rules in a rules-based standard permit the manager's preferred stand, the manager can use them to justify his/her preferred accounting treatment. Auditor-type is unlikely to matter with a rules-based standard. For example, a rules-oriented auditor will adhere to the application of the rules, and therefore agree with the manager's preferred accounting treatment. A client-oriented auditor is also unlikely to curtail managers' opportunistic accounting practices in the presence of rules-based standards. A client-oriented auditor's stand is influenced by what the manager prefers as long as it is not GAAP-inconsistent. In a rules-based regime, a client-oriented auditor is likely to permit interpretations of the rules and associated loopholes that favor the manager's preferred accounting treatment. Finally, a principles-oriented auditor may prefer not to adopt the client's proposed accounting treatment to the extent that application of the rules is contrary to the economic substance of the transaction or underlying accounting principles. However, the principles-oriented auditor's negotiation power is effectively constrained by the rules in the accounting standards, and the manager is likely to achieve his/her preferred accounting treatment. This discussion suggests the following hypothesis:

H1a: When accounting standards are rules-based, managers' likelihood of adopting an opportunistic accounting treatment will not differ as a function of whether the auditor is rules-oriented, client-oriented, or principles-oriented.

Just as GAAP-type can constrain the effects of auditor-type (as in the illustration above), the reverse is also true—the efficacy of a particular type of accounting standard can be constrained by auditor-type. This is the case with a principles-based standard. As argued by standard-setters and regulators, an accounting standard that is principles-based can *potentially* reduce opportunities for managers to achieve their incentive-driven accounting choices. However, this is contingent on auditor-type. If the auditor is more rules-oriented, then the underlying principle of the principles-based standard is not likely to be enforced, and managers can use seemingly rule-like provisions or exceptions in the standard to achieve their desired accounting choices, even if these are inconsistent with underlying accounting principles. More importantly, to the extent that managers engage in transaction structuring to get around the standard and achieve their preferred accounting treatment, the rules-oriented auditor is unlikely to object because he/she is less likely to attend to the underlying economics of the transaction or accounting principles. Similarly, a client-oriented auditor is sympathetic to the client, and will not object to the use of exceptions or transaction structuring even with a principles-based standard.

This discussion suggests that the presence of only a principles-oriented auditor (absent a principles-based standard) or a principles-based standard (absent a principles-oriented auditor) is unlikely to curtail aggressive accounting practices of managers. Rather, a combination of both a

principles-oriented auditor *and* a principles-based standard is necessary. In particular, a principles-based GAAP standard supports the principles-oriented auditor's stand and strengthens his/her negotiation power relative to that of the manager. Managers who are aware that they are dealing with a principles-oriented auditor in the context of a principles-based standard should therefore be less likely to attempt opportunistic interpretations of GAAP to achieve a preferred accounting treatment—the auditor is unlikely to support such GAAP interpretations, and a principles-based accounting standard offers little negotiation power to the manager.

Our hypothesis is formally stated below:

H1b: When accounting standards are principles-based, managers' likelihood of adopting an opportunistic accounting treatment will be lower when the auditor is principles-oriented than when the auditor is either rules-oriented or client-oriented.

III. RESEARCH METHOD

Participants

Participants are 90 financial managers in Canada, with mean business experience of 19.5 years. Sixty-three participants (70 percent) have current or prior experience as chief financial officers (mean experience as chief financial officer = 6.7 years), and 26 (29 percent) of them have worked as CFOs in a listed company. The remaining 27 participants are financial controllers/managers or equivalent with significant business experience (mean = 10.8 years). Several professional associations—Financial Executives International (FEI) national office in Toronto as well as FEI chapters in Calgary and Edmonton, Institute of Chartered Accountants of Alberta, and Certified Management Accountants (Alberta)—assisted the researchers in identifying potential participants. These organizations distributed email messages describing the research study to their members, advertised the study in their professional newsletters, and encouraged members to participate in the study. Executives of these organizations, and some prominent corporate directors who have national professional networks, also provided an endorsement of the study. One key inducement for financial managers to participate in the study was the promise of personalized feedback from a Big 4 audit firm partner who would read their memo.[5] One of the researchers then contacted 169 potential participants by phone or via email to seek their participation. Upon their consent, 163 research instruments were sent to participants by courier mail. Ninety completed instruments were returned to the researcher, for a response rate of 55 percent.[6]

Design

We employ a 2 × 3 between-subjects design, with GAAP-type (principles-based, rules-based) and Auditor-type (principles-oriented, rules-oriented, and client-oriented) as independent variables. For all treatments, participants are told to assume that they are the CFO of Oilsands, Inc., a publicly listed oil extraction company. The GAAP-type manipulation varies the accounting standard applicable to the client. In the rules-based condition, participants are told that Oilsands operates under Canadian GAAP. The manipulation instructs participants to assume that the client is "operating under current Canadian GAAP. HB 3065 proposes a series of very specific tests (e.g.,

[5] We recruited three Big 4 firm audit partners (two from firm 1 and one from firm 2) to provide feedback to participants (one client-oriented, one principles-oriented, and one rules-oriented). A partner in a Big 4 firm gave us his opinion about which audit partners were client-oriented, principles-oriented, and rules-oriented. We then called these audit partners and asked them to rate themselves as being either client-, principles-, or rules-oriented. If both our informant's judgment and the self-rating by the partner matched, then we asked the partner to provide feedback in that role.

[6] We sent out two separate calls for participation and administrations of the experiment. In the first (second) administration, 111 (52) research instruments were sent, and 61 (29) were returned.

present value of lease payments > 90% of fair market value of the asset; lease term > 75% of useful life of asset) to determine the type of lease involved (capital, operating) and its associated valuation and disclosure in the financial statements."

In the principles-based condition, participants are told that Oilsands operates under International Accounting Standards (IAS), and that under IAS, the new proposed lease standard based on a G4+1 research study (IAS 17) removes the four conditions in the current lease standard, and requires the present value of all leases to be capitalized and amortized (all leases are capital leases). Thus, our GAAP-type manipulation captures the key features of a principles- (versus rules-) based standard (SEC 2003). As opposed to the rules-based standard, the principles-based standard has a clearly stated accounting principle (leases are associated with corresponding assets and liabilities that should be recorded on the balance sheet), no bright-line tests, and no exceptions to the principles underlying the standard (although loopholes remain available for moving the transaction off balance sheet due to scope, valuation methods, and other exceptions in the existing IAS 17 standard).

The Auditor-type manipulation varies whether the audit partner in charge of the audit is known to be principles-oriented, rules-oriented, or client-oriented. The principles-oriented audit partner is described as "neutral and objective. He is an economic substance auditor who makes judgments based on what he assesses to be the economic substance of a transaction. In fact, on a couple of occasions, it has been said that, compared to the average partner, he is more concerned about capturing the substance of a transaction than about strictly following the Handbook." The rules-oriented audit partner is described as: "a very technical person. He is known to be a strictly by-the-Handbook auditor who follows Handbook rules very closely. In fact, on a couple of occasions, it has been said that, compared to the average partner, he is more concerned about strictly following the Handbook than about the economic substance of the transaction." Finally, the client-oriented audit partner is described as "entrepreneurial and client service-oriented, and open to client concerns. He is a client-oriented person who cares about the client's views. In fact, on a couple of occasions, it has been said that he grants the client a little more latitude than the average partner."

Procedure

The case was developed using disclosures, as provided by a large publicly listed oil company, of a significant off balance sheet transaction. The case materials contain background information along with a summary of selected income and balance sheet items related to Oilsands, a major oil extraction company in Canada.[7] Participants are told that Oilsands had earlier made an announcement about a major facility expansion. The chief executive officer indicates the need to first build a new power plant, and that his preference is for the power plant transaction to be structured "off balance sheet" as an operating lease rather than a capital lease. The reason is that Oilsands intends to seek additional debt financing for the major facility expansion later on, and the CEO is concerned that showing the power plant transaction as debt may jeopardize the company's ability to obtain additional debt financing.

We provided participants with details about the proposed leasing arrangement with a financial institution. The present value of the lease payments is 87.5 percent of the power plant's fair market value. The lease is for 25 years, which is about 71.4 percent of the power plant's useful life of 35 years. There is no ownership transfer at the end of the lease period, but there is an option to

[7] Participants in the first administration of the experiment also completed, in random order, a case relating to the underlying basis (e.g., based on direct hours, time spent, billings to date) to recognize revenue from a construction contract under Canadian GAAP. We find no main or interactive effect of order on our dependent variable when we include order as an additional independent variable. We find no effect of auditor-type for this construction contract case.

purchase the power plant at the end of the lease period at its fair market value. Oilsands incurs a penalty payment of 20 percent of the fair market value at the end of the lease if it does not do so. We make this design choice for two reasons. First, this setting keeps the accounting reporting flexibility constant (at a low level) for both rules-based and principles-based standards. For instance, the transaction meets the explicitly stated bright-line tests listed in the rules-based Canadian GAAP for classifying the transaction as an operating lease (present value of lease payments $<$ 90 percent of fair market value of the asset; lease term $<$ 75 percent of useful life of asset; no bargain purchase option; no ownership transfer). Hence, there is little reporting flexibility under Canadian (rules-based) GAAP going by the bright-line rules that point toward the lease being classified as an operating lease, consistent with management's incentives. Similarly, under the proposed IAS (principles-based) GAAP, there is also little reporting flexibility as the transaction will be required to be reported on the balance sheet, inconsistent with management's incentives. Second, because our focus is on the potential benefits of a principles-based standard (relative to the rules-based standard as a comparison), this creates a situation where there is conflict between the accounting treatment the manager prefers and that proposed by the principles-based standard. Such a setting enables us to assess whether auditor-type (specifically, a rules-oriented auditor) can undermine the benefits of a principles-based standard. Similarly, in the case of the rules-based standard where there is no conflict between managers' preference and the accounting treatment suggested by the rules, this setting enables us to assess whether auditor-type (specifically, a principles-oriented auditor) can overcome the rules and enforce the underlying economics of the transaction.

Participants also read a draft disclosure note prepared by the accounting department on the lease transaction, classified as an operating lease. They were told that the CEO has reviewed the draft, and felt that the draft note drew too much attention to the power plant transaction. After reading the case, participants were required to indicate how they would set up the power plant transaction, the basis for their recommendations, and how Oilsands should disclose the transaction in the financial statements. They were also asked to write a memo to convince the auditor on the propriety of the accounting treatment of this lease.

Our primary dependent variable relates to whether participants seek to take the transaction off balance sheet (that is, not as a capital lease) in accordance with management's incentives. This dependent variable is important because regulators have been concerned about material off balance sheet transactions and management's attempts at transaction structuring (see interview with Sir David Tweedie, as reported in Sawers 2008). To measure this, we asked an open-ended question related to how participants suggest the power plant transaction be set up, given the CEO's concern.[8] Two doctoral students who are chartered accountants independently coded the responses into three categories: operating lease, transaction structuring to take the transaction off balance sheet, and capital lease. There was 100 percent agreement. The coders were not aware of the experimental conditions associated with the memos. For the purpose of our dependent variable (whether transaction is taken off balance sheet), the first two categories were coded as off balance sheet transactions.

IV. RESULTS

Manipulation Checks

As a check on the GAAP manipulation, we asked participants whether the company operates under Canadian, U.S., or IASB GAAP. Four participants answered the question incorrectly. The

[8] We asked an open-ended question to allow participants flexibility in structuring the transaction (including non-lease transaction). Asking a close-ended question with pre-specified options would limit their responses to the options provided, and also provide a cue that non-lease transactions are possible.

manipulation check question related to auditor-type asks participants to indicate whether the audit partner can best be characterized as strict (Mr. GAAP), neutral and objective, or entrepreneurial and client service-oriented. Five participants answered this question incorrectly. We obtain the same results after excluding these participants who failed the manipulation check.

Test of Hypothesis

Table 1, Panel A presents the descriptive statistics for the proportion of participants who opted to record the transaction off balance sheet; that is, not as a capital lease. Table 1, Panels B and C show the categorical ANOVA results along with the planned contrasts. Hypothesis 1a predicts that with rules-based standards, auditor-type does not matter in terms of participants' taking the transaction off balance sheet, while H1b predicts that auditor-type matters with principles-based standards. This suggests an interaction between GAAP-type and Auditor-type. As shown in Table 1, Panel B, we find a statistically significant interaction between GAAP-type and Auditor-type ($p = 0.010$).[9]

As Table 1, Panel C shows, when standards are rules-based, we find no simple main effect of auditor-type ($p = 0.423$), and all possible pairwise comparisons are not statistically significant (all $p > 0.208$). These results are consistent with H1a. In contrast, H1b predicts with principles-based standards, auditor-type matters in that the proportion of financial managers taking the transaction off balance sheet is lower when the auditor is principles-oriented than when the auditor is either rules-oriented or client-oriented. Results show that, consistent with H1b, when standards are principles-based, there is a simple main effect of auditor-type ($p = 0.010$). Pairwise comparisons show that, as predicted, the proportion of participants taking the transaction off balance sheet is lower with a principles-oriented auditor than with either a rules-oriented auditor ($p = 0.002$; one-tailed) or client-oriented auditor ($p = 0.036$; one-tailed). The difference in mean proportion when the auditor is rules-oriented or client-oriented is not statistically significant ($p = 0.113$).

A corollary of H1a and H1b is that a combination of both a principles-oriented auditor *and* a principles-based standard is necessary to curb managers' aggressive accounting practices. This suggests that the proportion of participants taking the transaction off balance sheet is likely to be lowest in the principles-based standard/principles-oriented auditor condition. A specific contrast test comparing the mean proportion in the principles-based standard/principles-oriented auditor condition compared to the average in the other five conditions is statistically significant ($p = 0.001$, one-tailed). We also made pairwise comparisons between the principles-based standard/principles-oriented auditor condition and each of the other five conditions, and find that the proportion is always significantly lower in the principles-based standard/principles-oriented auditor condition than in the other conditions (all $p < 0.036$, one-tailed).[10]

Transaction Structuring

We analyze the proportion of participants who opt to structure their transactions to achieve the objective of having the transaction taken off the balance sheet. This is a costly option as it involves the firm incurring costs to restructure their original planned transaction to avoid having the transaction being recognized on the balance sheet. Participants who engaged in transaction restructur-

[9] To assess whether our results vary by the two administrations of the experiment, we test whether the GAAP-type × Auditor-type interaction predicted in H1 holds when we include Administration (first, second) as a third independent variable. Results of a categorical ANOVA with proportion of participants opting to take the transaction off balance sheet shows that the GAAP-type × Auditor-type interaction continues to be significant ($p = 0.008$). There is also a marginal GAAP-type × Auditor-type × Administration interaction ($p = 0.095$)—we find that the two-way GAAP-type × Auditor-type is always significant in both administrations, but is more significant in Administration 1 ($p = 0.000$) and less so in Administration 2 ($p = 0.025$).

[10] An overall test assessing for differences among any of these five conditions is not statistically significant ($p = 0.339$).

TABLE 1

Proportion of Managers Selecting Incentive-Consistent Off Balance Sheet Disclosure

Panel A: Descriptive Statistics (percentage, sample size)

	Auditor-Type			
	Principles-Oriented	**Client-Oriented**	**Rules-Oriented**	**Overall**
GAAP-type				
Rules-based	86.7%	80.0%	66.7%	77.8%
	n = 15	n = 15	n = 15	n = 45
Principles-based	26.7%	60.0%	86.7%	57.8%
	n = 15	n = 15	n = 15	n = 45
Overall	56.7%	70.0%	76.7%	67.8%
	n = 30	n = 30	n = 30	n = 90

Panel B: Results of Categorical ANOVA for Lease Treatment Choice

	df	**Chi-Square**	**p-value**
Intercept	1	10.93	0.001
Auditor-type	2	1.62	0.444
GAAP-type	1	2.90	0.089
Auditor-type × GAAP-type	2	9.22	0.010

Panel C: Planned Comparisons

Planned Mean Comparison	χ^2	**p-value**
Test of H1a (effects of auditor-type when accounting standards are rules-based)	1.72	0.423
Pairwise comparisons of auditor-type (when accounting standards are rules-based)		
$\%_{\text{Principles–Oriented Auditor; Rules–Based Standard}}$ versus $\%_{\text{Rules–Oriented Auditor; Rules–Based Standard}}$	1.58	0.208
$\%_{\text{Principles–Oriented Auditor; Rules–Based Standard}}$ versus $\%_{\text{Client–Oriented Auditor; Rules–Based Standard}}$	0.24	0.626

(continued on next page)

Panel C: Planned Comparisons

Planned Mean Comparison	χ^2	p-value
%Rules–Oriented Auditor; Rules–Based Standard versus %Client–Oriented Auditor; Rules–Based Standard	0.67	0.413
Test of H1b (effects of auditor-type when accounting standards are principles-based)	9.30	0.010
Pairwise comparisons of auditor-type (when accounting standards are principles-based)		
%Principles–Oriented Auditor; Principles–Based Standard versus %Rules–Oriented Auditor; Principles–Based Standard	9.06	0.002*
%Principles–Oriented Auditor; Principles–Based Standard versus %Client–Oriented Auditor; Principles–Based Standard	3.25	0.036*
%Rules–Oriented Auditor; Principles–Based Standard versus %Client–Oriented Auditor; Principles–Based Standard	2.52	0.113

* One-tailed test.

ing proposed a variety of alternative structures. The simplest was to get a utility to construct and run a power plant and then sign a long-term contract to purchase the power (an off balance sheet commitment). More elaborate structures involved setting up a trust, or getting together with other oil companies to set up a variable interest entity.

Following our theory and experimental design, when the standard is rules-based, there is little need for participants to engage in costly transaction restructuring regardless of auditor-type because the rules permit the classification of the transaction as a lease. In contrast, when the standard is principles-based and requires the transaction to be reported on the balance sheet, transaction restructuring is expected to be higher because this is one of the key approaches to keep the transaction off balance sheet. However, this outcome is dependent on auditor-type. Even with a principles-based standard, transaction restructuring is expected to be low when the auditor is principles-oriented because participants likely anticipate the auditor would still require the restructured transaction to be recognized in the balance sheet. The rules-oriented or client-oriented auditor is less likely to insist on this.

Descriptive statistics are shown in Table 2, Panel A. Table 2, Panel B shows the results of the categorical ANOVA. We find a statistically significant main effect of GAAP ($p = 0.021$), but no significant main effect of Auditor-type ($p = 0.561$) or interaction effect ($p = 0.213$). Table 2, Panel C reports the simple effects of auditor-type when the standard is either rules- or principles-based.

Overall, only 23.3 percent of the participants opt for transaction structuring. When the standard is rules-based, a very low proportion of participants engage in transaction restructuring (11.1 percent). Consistent with our expectation, there is no significant simple main effect of auditor-type (all $p = 0.805$), and all pairwise comparisons within auditor-type are not statistically significant ($p > 0.550$). As predicted, when the standard is principles-based, we find a marginally significant simple main effect of auditor-type ($p = 0.092$). With a principles-based standard, the mean proportion of participants undertaking transaction structuring when the auditor is principles-oriented is very low (13.3 percent), and is significantly lower than when the auditor is client-oriented ($p = 0.056$; one-tailed) or rules-oriented ($p = 0.015$; one-tailed), where much of the transaction structuring occurs (40.0 percent and 53.3 percent, respectively). A specific contrast test comparing the mean proportion in these latter two conditions versus the other four conditions is statistically significant ($p = 0.001$; one-tailed).

Thought Units

Our theoretical reasoning implies that accounting standards (principles- or rules-based) serve as a source of negotiation power in that they help justify a particular accounting choice. However, the need for and ease of justification likely vary with the preferred outcome of the manager. Managers who wish to record the transaction as a capital lease should find it easier to justify their preference when the accounting standard is principles-based. Likewise, managers who prefer to record the transaction off balance sheet should find it easier to justify their preference when the accounting standard is rules-based. This suggests an accounting choice by GAAP-type interaction. This interaction may be weaker when participants in the principles-based/on balance sheet (rules-based/off balance sheet) condition anticipate interacting with a principles-oriented (rules-oriented) auditor because they believe such an auditor is already convinced and there is, therefore, less need to make a strong justification. On the other hand, there may be no moderating effect of auditor-type if participants perceive that all auditors expect to see appropriate justifications for managers' accounting choices.

To test this prediction, we first content code participants' rationale memos to identify the distinct thought units documented in the memos that participants used to persuade their assigned audit partner regarding the propriety of the proposed accounting treatment. We develop a coding

TABLE 2

Proportion of Managers Opting for Transaction Structuring

Panel A: Descriptive Statistics (mean, standard deviation, sample size)

	Auditor-Type			
	Principles-Oriented	Client-Oriented	Rules-Oriented	Overall
GAAP-type				
Rules-based	13.3%	13.3%	6.7%	11.11%
	n = 15	n = 15	n = 15	n = 45
Principles-based	13.3%	40.0%	53.3%	35.6%
	n = 15	n = 15	n = 15	n = 45
Overall	13.3%	26.7%	30.0%	23.3%
	n = 30	n = 30	n = 30	n = 90

Panel B: Results of Categorical ANOVA for Lease Treatment Choice

	df	Chi-Square	p-value
Intercept	1	21.72	0.000
Auditor-type	2	1.16	0.561
GAAP-type	1	5.37	0.021
Auditor-type × GAAP-type	2	3.11	0.213

Panel C: Planned Comparisons

Planned Mean Comparison	χ^2	p-value
Test of H1a (effects of auditor-type when accounting standards are rules-based)	0.43	0.805
Pairwise comparisons of auditor-type (when accounting standards are rules-based)		
$\%_{\text{Principles–Oriented Auditor; Rules–Based Standard}}$ versus $\%_{\text{Rules–Oriented Auditor; Rules–Based Standard}}$	0.36	0.550
$\%_{\text{Principles–Oriented Auditor; Rules–Based Standard}}$ versus $\%_{\text{Client–Oriented Auditor; Rules–Based Standard}}$	0.00	1.000

(continued on next page)

Panel C: Planned Comparisons

Planned Mean Comparison	χ^2	p-value
$\%_{\text{Rules–Oriented Auditor; Rules–Based Standard}}$ versus $\%_{\text{Client–Oriented Auditor; Rules–Based Standard}}$	0.36	0.550
Test of H1b (effects of auditor-type when accounting standards are principles-based)	4.78	0.092
Pairwise comparisons of auditor-type (when accounting standards are principles-based)		
$\%_{\text{Principles–Oriented Auditor; Principles–Based Standard}}$ versus $\%_{\text{Rules–Oriented Auditor; Principles–Based Standard}}$	4.76	0.015*
$\%_{\text{Principles–Oriented Auditor; Principles–Based Standard}}$ versus $\%_{\text{Client–Oriented Auditor; Principles–Based Standard}}$	2.52	0.056*
$\%_{\text{Rules–Oriented Auditor; Principles–Based Standard}}$ versus $\%_{\text{Client–Oriented Auditor; Principles–Based Standard}}$	0.53	0.466

* One-tailed test.

scheme using Koonce et al. (1995) as a guide. The same two coders (doctoral students who are chartered accountants) independently coded the memos written by participants. Inter-rater agreement was 89 percent, and discrepancies were resolved by discussion. The coders were not aware of the experimental conditions associated with the memos.

Overall, eight different categories of thought units were identified: (1) specific arguments on whether the lease is/is not capital; (2) negotiating the terms of the lease; (3) setting up a trust or special purpose entity; (4) nature of the business and whether it falls outside scope of standard; (5) details of nature of disclosure after selection of classification; (6) recapitulation of case facts without adding new insights; (7) greetings and pleasantries unrelated to case facts, and (8) other thoughts. We find no main or interaction effect of Auditor-type or GAAP-type on the total number of thought units (smallest $p = 0.296$).

We compute the proportion of relevant thought units as a proportion of total thought units ($Proportion_{relevant}$) as the dependent variable. The first four categories ((1) to (4)) are arguments directly relevant to justifying the proposed classification by the manager. The remaining four ((5) to (8)) are less relevant for convincing the auditor to agree with the accounting measurement proposed by the manager. For example, the thought units on disclosure relate to discussion on the disclosure of lease terms, but not the rationale for classifying the lease as operating or capital. These disclosures *per se* would not convince the auditor to allow this measurement treatment. We interpret a higher $Proportion_{relevant}$ to mean that the participant is better able to garner support or justification for his/her proposed classification.

We conduct an ANOVA with Auditor-type, GAAP-type, and Accounting Choice (operating versus capital lease) as independent variables.[11] Table 3, Panel A reports the proportion of relevant thought units documented for all managers, while Table 3, Panels B and C report the corresponding proportions when the managers select a capital lease option and report off balance sheet, respectively. The ANOVA results are shown in Table 3, Panel D. We find a significant GAAP-type by Accounting Choice interaction ($p = 0.001$), and the p-values are not statistically significant either for the other main effects (smallest $p = 0.153$) or interaction effects (smallest $p = 0.332$). Further analysis indicates that when participants make a capital lease option (Table 3, Panel B), $Proportion_{relevant}$ is higher for a principles-based GAAP than a rules-based GAAP (mean = 0.801 and 0.470, respectively; $p = 0.005$; one-tailed). We find the opposite for participants who select an off balance sheet treatment (Table 3, Panel C); $Proportion_{relevant}$ is marginally lower for principles-based GAAP than rules-based GAAP (mean = 0.475 and 0.602, respectively; $p = 0.030$; one-tailed).[12]

These results are consistent with the notion that principles- versus rules-based standards serve as sources of negotiation power insofar as they influence the extent that financial managers can justify their accounting choices in their negotiations with the auditor.

V. CONCLUSION

This study investigates the effect of principles-based versus rules-based accounting standards on the reporting positions of managers who interact with three different auditor-types (an auditor who is principles-oriented, rules-oriented, or client-oriented). Our experiment comparing the ef-

[11] When we run an overall Auditor-type by GAAP-type ANOVA (without partitioning by accounting choice), we find no significant main (smallest $p = 0.498$) or interaction effect ($p = 0.135$). One participant did not write the memo.

[12] In fact, when participants select an off balance sheet transaction, a contrast test indicates that Proportion relevant is lowest in the Principles-based standard/Principles-oriented auditor condition (mean = 0.262) than that in the other five conditions (mean = 0.570; $p = 0.023$; one-tailed). This finding is consistent with the argument that the Principles-based standard/Principles-oriented auditor condition constrains managers' negotiation power (and justifications they can come up with) when they attempt to make a case for taking the transaction off balance sheet. A caveat to this analysis is that the sample size in the Principles-based standard/Principles-oriented condition is very small.

TABLE 3

Proportion of Relevant Thought Units in Managers' Rationale Memos

Panel A: Descriptive Statistics for Overall Sample (proportion, sample size)

	Auditor-Type			
	Principles-Oriented	Client-Oriented	Rules-Oriented	Overall
GAAP-type				
Rules-based	0.581	0.530	0.607	0.573
	n = 15	n = 15	n = 15	n = 45
Principles-based	0.654	0.718	0.465	0.616
	n = 15	n = 15	n = 14	n = 44
Overall	0.618	0.624	0.539	0.594
	n = 30	n = 30	n = 29	n = 89

Panel B: Descriptive Statistics for Managers Who Select Capital Lease Option (proportion, sample size)

	Auditor-Type			
	Principles-Oriented	Client-Oriented	Rules-Oriented	Overall
GAAP-type				
Rules-based	0.303	0.397	0.580	0.470
	n = 2	n = 3	n = 5	n = 10
Principles-based	0.797	0.795	0.844	0.801
	n = 11	n = 6	n = 2	n = 19
Overall	0.721	0.663	0.656	0.687
	n = 13	n = 9	n = 7	n = 29

(continued on next page)

Panel C: Descriptive Statistics for Managers Who Select Off Balance Sheet Option (proportion, sample size)

	Auditor-Type			
	Principles-Oriented	**Client-Oriented**	**Rules-Oriented**	**Overall**
GAAP-type				
Rules-based	0.624	0.563	0.620	0.602
	n = 13	n = 12	n = 10	n = 35
Principles-based	0.262	0.666	0.402	0.475
	n = 4	n = 9	n = 12	n = 25
Overall	0.539	0.607	0.501	0.549
	n = 17	n = 21	n = 22	n = 60

Panel D: ANOVA Results

Source	**Sum of Squares**	**df**	**F**	**p-value**
Auditor-type	0.153	2	0.865	0.425
GAAP-type	0.185	1	2.083	0.153
Accounting choice	0.134	1	1.514	0.222
Auditor-type * GAAP-type	0.153	2	0.862	0.426
Auditor-type * Accounting choice	0.124	2	0.701	0.499
GAAP-type * Accounting choice	1.069	1	12.064	0.001
Auditor-type * GAAP-type * Accounting choice	0.198	2	1.119	0.332
Error	6.822	77		

fects of a proposed standard change and a current rules-based GAAP standard provides useful *ex ante* evidence for standard-setters who are considering adoption of principles-based standards.

Our results indicate that with a rules-based standard, auditor-type does not influence managers' propensity to take the proposed lease transaction off balance sheet. An implication of this result is that even when the auditor is principles-oriented and seeks to enforce the economic substance of a transaction, he/she is constrained from doing so in the presence of a rules-based standard. We find that with a principles-based standard, the auditor-type matters. Specifically, the propensity to take the proposed lease transaction off balance sheet is lowest when the auditor is principles-oriented compared to the situation when the auditor is rules-oriented or client-oriented. In fact, a combination of a principles-oriented auditor and a principles-based standard results in the lowest incidence of the proposed lease transaction being reported off balance sheet relative to any other combination of auditor-type and standard-type. These results are consistent with the testimony of Sir David Tweedie (2002), in which he argues that a principles-based accounting system requires both preparers and auditors to work together to reflect the economic substance of transactions in the financial statements. Thus, managers may still engage in aggressive reporting when the principles-oriented auditor operates under rules-based standards, or when a principles-based standard is enforced by rules-oriented or client-oriented auditors.

Our analysis of participants' transaction restructuring attempts indicates a similar picture. They are less likely to engage in costly restructuring attempts when the standard is rules-based (as there is no need to) or when the standard is principles-based and auditor is principles-oriented (there is a need to but the auditor will object). Transaction restructuring attempts are more prevalent when the standard is principles-based (necessitating costly restructuring) and the auditor will be supportive (when s/he is rules-oriented or client-oriented). Thus, potential benefits of a principles-based standard can be obfuscated by auditors who are rules- and/or client-oriented.

Our results with respect to financial managers' propensity to restructure the transaction also add to extant auditor-client negotiation literature. The results suggest that the auditor's bargaining power relative to the client (vested through the type of accounting standards) has effects on pre-negotiation transactions undertaken by the client, not only at the time of negotiation as has been commonly looked at in the literature. Furthermore, the subject matter brought to the table for auditor-client negotiation may have been the outcome of a strategic decision made by the financial manager in response to the type of auditor who will engage in the negotiation.

The principles-based standard that we examine is modeled after a proposal by the G4+1 countries to report all leases on the balance sheet, which is also consistent with the current proposal by the FASB and IASB. Our results suggest that the benefit of the proposed lease accounting standard may only be realized if we *also* have an auditor who focuses on reflecting the substance of the transaction in the financial statements. A primary message of our paper is that a better reporting environment is created only if both the accounting standards and the auditor are principles-oriented. Changing only one of these variables (accounting rules or the auditor) is less likely to be effective. For instance, when the standard is rules-based, a principles-oriented auditor's ability to deter aggressive accounting practices is limited by the rules in the standard; in contrast, when the standard is principles-based, the principles-oriented auditor is better able to do so (relative to a rules-oriented or client-oriented auditor). Thus, proposals for changing accounting standards require a simultaneous consideration of changes in the mindset of the auditor as well. Furthermore, to the extent that auditors are currently not sufficiently principles-oriented, the need for training auditors to be so becomes increasingly important (SEC 2003).

One limitation of our study is that the accounting standard we examine is restricted to lease accounting, which may not be representative of all accounting standards. In addition, we test a proposed lease guidance, and not an existing lease standard. It is possible that the manner by which managers react to a proposed standard may be different from how they react to an actual

standard. For instance, once they develop familiarity with such a standard, they may develop new ways to get around the standard. Also, CICA HB 3065 and the IAS 17-plus guidance likely differ in other aspects (e.g., wording, format), so while we can attribute our results to GAAP-type differences, we cannot isolate the specific aspect leading to our results. Another limitation is that auditor-type (especially client-oriented auditors) might matter more when managers prefer outcomes that are inconsistent with both the letter and intent of the standards. Furthermore, we do not determine how auditors respond to principles-based versus rules-based standards, and the extent to which auditors are currently principles-oriented versus rule-oriented. In our study, the auditor does not negotiate with the financial manager. Future research can investigate these issues. Our study provides a starting point to understand how it is necessary to investigate *both* the characteristics of accounting standards and auditors to understand reporting decisions of managers.

REFERENCES

Antle, R., and B. Nalebuff. 1991. Conservatism and auditor-client negotiations. *Journal of Accounting Research* 29 (Supplement): 31–54.

Bacharach, S., and E. Lawler. 1981. *Bargaining Power, Tactics, Outcomes*. San Francisco, CA: Jossey-Bass Inc.

Cialdini, R. 2007. Descriptive social norms as underappreciated sources of social control. *Psychometrika* 72 (2): 263–268.

Coleman, J. 1990. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Essaides, N. 2006. Fear factor: Treasury accounting. May 25. Available at: http://www.iTreasurer.com.

Financial Accounting Standards Board (FASB). 2004. *Response to SEC Study on the Adoption of A Principles-Based Accounting System*. Norwalk, CT: FASB.

Financial Accounting Standards Board/International Accounting Standards Board (FASB/IASB). 2009. *Leases: Preliminary Views*. Norwalk, CT: FASB.

Fiske, S. T., and J. L. Berdahl. 2007. Social power. In *Social Psychology: A Handbook of Basic Principles*, edited by Kruglanski, A., and E. T. Higgins. New York, NY: The Guilford Press.

French, R. P. Jr. and B. Raven. 1959. The bases of social power. In *Studies in Social Power*, edited by D. Cartwright. Ann Arbor, MI: University of Michigan, Institute for Social Research.

Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (3): 535–563.

———, S. McCracken, and S. Salterio. 2005. Negotiations over accounting issues: The congruency of audit partner and chief financial officer recalls. *Auditing: A Journal of Practice & Theory* 24 (Supplement): 171–193.

Grout, P., I. Jewitt, C. Pong, and G. Whittington. 1994. Áuditor professional judgment: Implications for regulation and the law. *Economic Policy* 9 (October): 308–351.

Hackenbrack, K., and M. W. Nelson. 1996. Auditors' incentives and their application of financial accounting standards. *The Accounting Review* 71 (1): 43–60.

Koonce, L., U. Anderson, and G. Marchant. 1995. Justification of decisions in auditing. *Journal of Accounting Research* 33 (2): 369–384.

Lim, C. Y., and H. T. Tan. 2008. Non-audit service fees and audit quality: The impact of audit specialization. *Journal of Accounting Research* 46 (1): 199–246.

Magee, J. C., A. D. Galinsky, and D. H. Gruenfeld. 2007. Power, propensity to negotiate, and moving first in competitive interactions. *Personality and Social Psychology Bulletin* 33 (2): 200–212.

Millman, G. J. 2005. Companies sensing a big chill with auditors. *Financial Executive* 21 (June): 22–25.

Moore, D. A., P. E. Tetlock, L. Tanlu, and M. H. Bazerman. 2006. Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review* 31 (1): 10–29.

Nailor, H., and A. Lennard. 2000. *Leases: Implementation of A New Approach*. FASB-Financial Accounting Series Special Report. Norwalk, CT: FASB.

Pruitt, D., and P. Carnevale. 1993. *Negotiation in Social Conflict*. Buckingham, UK: Open University Press.

Public Company Accounting Oversight Board (PCAOB). 2007. *Observations of Auditors' Implementation of PCAOB Standards Relating to Auditors' Responsibilities with Respect to Fraud.* Release No. 2007-01. Washington, D.C.: PCAOB.

Ryan, S., R. H. Herz, T. E. Iannaconi, L. A. Maines, K. G. Palepu, K. Schipper, C. M. Schrand, D. J. Skinner, and L. Vincent. 2001. Evaluation of the lease accounting proposed in G4+1 Special Report. *Accounting Horizons* 15 (3): 289–298.

Sanchez, M. H., C. P. Agoglia, and R. C. Hatfield. 2007. The effect of auditors' use of a reciprocity-based strategy on auditor-client negotiations. *The Accounting Review* 82 (1): 205–240.

Sawers, A. 2008. Profile: Sir David Tweedie, Setting the standards. *Accountancy Age* (April 17).

Scott, H. S. 2006. *Interim Report of the Committee on Capital Markets Regulation.* November 30. Available at http://www.capmktsreg.org.

Securities and Exchange Commission (SEC). 2003. *Study Pursuant to Section 108(d) of the Sarbanes-Oxley Act of 2002 on Adoption by the United States Financial Reporting System of a Principles-Based Accounting System.* Submitted to Committee on Banking, Housing and Urban Affairs of the U.S. Senate and Committee on Financial Services of the U.S. House of Representativeness. Washington, D.C.: Government Printing Office.

———. 2008. *Progress Report of the Advisory Committee on Improvements to Financial Reporting to the United States Securities and Exchange Commission.* Washington, D.C.: Government Printing Office.

Tan, H. T., and K. Trotman. 2010. Effects of auditors' concession timing on financial officers' negotiation judgments. *Contemporary Accounting Research.*

Taub, S.. 2005. Wary auditors, clients worry FASB's Herz. *CFO.com* (January 25).

Tweedie, D. 2002. Testimony before the U.S. Senate Committee on Banking, Housing and Urban Affairs. February 14. Washington, D.C.: Government Printing Office.

——— 2007. Keep it simple stupid! Can global standards be principle based? Speech made at the 2007 Ken Spencer Memorial Lecture. Available at: http://www.frc.gov.au/reports/other/Ken_Spencer_2007.asp.

Weirich, T. R., and R. W. Rouse. 2007. New guidance on SOX 404. *Journal of Corporate Accounting & Finance* 18 (4): 37–41.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1347–1374

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1347

Intangible Returns, Accruals, and Return Reversal: A Multiperiod Examination of the Accrual Anomaly

Robert J. Resutek
Dartmouth College

ABSTRACT: Prior studies employ a two-period empirical model and interpret the negative association between accruals in period one and returns in period two as evidence that investors misprice the information contained in accruals. In contrast to prior studies, I employ a three-period log-linear model decomposed from a firm's book-to-market ratio and show that investors do not misprice the information contained in accruals. My study shows that in the four-year period prior to accrual recognition, equity prices tend to be driven disproportionately by intangible returns, or returns not explained by accounting measures. Accordingly, the relation between prior period intangible returns and future period returns subsumes the relation between current period accruals and future returns. In addition, I link the accrual anomaly and the value/growth anomaly to a common economic mechanism (intangible returns) and show that a strong negative relation between external financing activities and future returns is not subsumed by the accrual anomaly.

Keywords: *accrual anomaly; book-to-market; external financing; investment; market efficiency.*

Data Availability: *The data used are available from the sources noted in the study.*

I. INTRODUCTION

Beginning with Sloan (1996), the accounting literature has produced an extensive stream of literature examining the negative relation between current period accruals and future period returns. From this stream of literature two primary explanations for the negative relation (the accrual anomaly) have emerged. The "persistence" explanation builds off Sloan (1996) and explains the accrual anomaly as caused by investor inability to distinguish the pricing

This paper is based on my dissertation at the University of Texas at Austin. I would like to thank the members of my committee: Robert Freeman (chair), Michael Clement, Ross Jennings, Sheridan Titman, and Yong Yu for their helpful comments and suggestions on earlier drafts of this paper. In addition I thank Rich Frankel, Kenneth French, Bill Kinney, Jonathan Lewellen, William Mayew, Richard Sansing, two anonymous reviewers, Mark Trombley (editor), and seminar participants at Washington University, Dartmouth, Stanford, Michigan State University and participants of the McCombs School of Business summer brown bag series. I am grateful for the financial support of the Deloitte & Touche Foundation, McCombs School of Business, and Tuck School of Business at Dartmouth.

Editor's note: Accepted by Mark Trombley.

Submitted: July 2008
Accepted: January 2010
Published Online: June 2010

implications of the differential persistence of the accrual and cash flow components of current earnings. The "growth" explanation builds off Fairfield et al. (2003) and suggests that investors misprice the growth information contained in accruals.

While these two streams of literature are distinct, they share some similarities. First, both explanations suggest that investors may be subject to behavioral biases that prevent them from properly incorporating publicly available accounting information into stock prices.[1,2] Second, as noted by Fairfield et al. (2003) and Zhang (2007), large positive accruals tend to occur during periods of growth and equity issuance. There is a strong negative relation between equity issuance activity and future returns (Loughran and Ritter 1995). Despite the empirical relation between equity issuance, accruals, and future returns, accrual studies in both streams have been relatively silent as to the effect that the interaction of these three variables on inferences made between future period returns and current period accruals.[3] Finally, both streams examine the relation between accruals and future returns primarily within the confines of a strict two-period model.[4] For example, studies in these streams generally examine whether variation in current period accruals are associated with variation in future returns after controlling for current period characteristics and factors known to be associated with future returns.

This study takes a different approach from that taken in prior accrual studies and finds results suggestive of a third explanation. Prior research has viewed a firm's book-to-market ratio (B/M) as a characteristic distinct from the firm's current period accruals. However, this view overlooks the fact that current period accruals affect B/M. Consider that B/M is comprised of four primary elements: (1) retained earnings, (2) invested equity, (3) expected cash flows, and (4) expected returns. Accordingly, changes in each of these elements explain changes in B/M. Given that accruals explain part of the changes in retained earnings, and that changes in each of the other three elements are associated with future returns, this study examines whether the evolution of B/M can shed new light on the negative association between current period accruals and future period returns.

The empirical design employed in this study builds on the work of Daniel and Titman (2006; hereafter, DT). DT note that the information investors incorporate into price can be partitioned into two parts: tangible information and intangible information. By assuming that a firm's market return over a given period is a good proxy for total information, DT classify the portion of a firm's realized market return that is explained by summary accounting measures as tangible information, or tangible returns. Accordingly, the residual realized return variation that is not explained by summary accounting measures is termed intangible information, or intangible returns. DT show in a two-period empirical model that the strong negative relation between price-scaled accounting variables (e.g., book-to-market, earnings-to-price) and future returns is not associated with total realized growth as measured by accounting variables. Rather, DT suggest that the return reversal associated with price-scaled accounting variables is driven by realized intangible returns.

I expand DT's empirical model into a three-period model to investigate the relation between prior period intangible returns in period one, accruals in period two, and future returns in period

---

[1] Assuming certain conditions relating to changes in a firm's expected returns and optimal investment in response to these changes, the growth explanation does not necessarily imply investor irrationality.

[2] Recent work by Mashruwala et al. (2006) and Khan (2008) suggest rational asset-pricing explanations of the accrual anomaly. Specifically, Mashruwala et al. suggest that arbitrageurs are prevented from capitalizing on the mispricing associated with accruals due to high arbitrage risk, while Khan proposes a four-factor ICAPM-based pricing model that explains the cross-sectional future return variation attributable to the current period accruals.

[3] Recent work by Cohen and Lys (2006) and Dechow et al. (2008) suggest that the negative relation between future returns and accruals subsumes the negative relation between future returns and external financing activities.

[4] Kothari et al. (2006) and Zhang (2007) consider the relation between accruals in the current year and firm growth in prior years, the current year, and subsequent years.

three. Intangible returns can generally be interpreted as capturing future growth options that have not been recognized by accounting performance measures. Given the discrete, conservative nature of financial accounting, intangible returns will tend to lead tangible returns as the market prices information more quickly than GAAP recognizes value. By examining the relation between current period accruals and prior period intangible returns, derived from the book-to-market ratio, I provide insight into a possible common link between the value/growth anomaly and the accrual anomaly.

Consistent with prior studies, I show that in two-period settings, the accrual component of log book returns explains a significant portion of future return variation. Controlling for firm growth captured by accounting measures of growth over the four-year period prior to accrual recognition, I show that the accrual component of log book returns still explains significant future return variation. However, once I control for prior period intangible returns (PPIR), I show that the return reversal previously attributed to accruals is subsumed by PPIR. More germane to the accounting literature, I show that once intangible returns from period one are controlled, the negative relation between accrual-based returns in period two and future returns in period three is insignificant. This pattern holds for multiple definitions and decompositions of accruals.

This study makes several contributions to the existing literature examining the relation between future returns and accounting measures of performance such as accruals. First, I provide empirical evidence that investors correctly incorporate the information contained in accruals into price. This evidence is significant given the widely held view that investors misprice the growth information contained in accruals (Fairfield et al. 2003; Zhang 2007) and/or the persistence characteristics of accruals (Sloan 1996; Richardson et al. 2005).

Second, this study contributes to research examining the relation between external financing activities, accruals, and future returns. Recent studies examining the relation between these three variables are mixed. Cohen and Lys (2006) and Dechow et al. (2008) suggest that the negative relation between future returns and external financing activities is essentially subsumed by the accrual anomaly, while Daniel and Titman (2006) show that their composite issuance variable subsumes any relation between future returns and accounting-based performance measures. The empirical evidence reported in this study serves to link these two disparate findings. Specifically, this study highlights an economically significant negative relation between future returns and equity issuance/repurchase activity and intangible returns while finding no significant association between future returns and accruals.

Third, this study employs an empirical design that addresses some of the concerns raised by Kraft et al. (2006, 2007) relating to accrual anomaly studies. Specifically, Kraft et al. (2007) show how the exclusion of variables correlated with accruals can affect inferences with future returns from Mishkin (1983) pricing tests. The empirical design used in this study endogenizes the pricing of all prior and current period information. This empirical design structure is important for two reasons. First, it mitigates the concern that significant relations between explanatory variables and future returns are caused by the omission of relevant variables. Second, the inferences drawn between future period returns and current period accruals are not contingent on the rational pricing of all information. Despite the agnostic stance on rational pricing this empirical design allows, the design still allows for inferences to be drawn about whether accruals are properly incorporated into price.

Fourth, this study provides a link between the value/growth anomaly and the accrual anomaly. Although the theories used to explain both anomalies are similar, these anomalies have generally been considered distinct. This study extends the idea of Beaver (2002) and Desai et al. (2004) who suggest that the accrual anomaly may be a function of the value/growth anomaly, providing the economic intuition and empirical evidence that shows these two anomalies are likely driven by the same underlying construct—specifically, intangible returns.

Section II introduces the empirical design and the economic intuition supporting it. Section III provides empirical evidence examining the relation between PPIR, current period accruals, and future returns. Sections IV and V discuss empirical evidence from additional robustness tests. Section VI provides concluding remarks.

II. RESEARCH DESIGN—METHODS AND INTUITION

Research Motivation and Variable Intuition

The value/growth anomaly is characterized by the positive association between future returns and price-scaled accounting variables such as book-to-market and earning-to-price. One of the more prominent explanations for the value/growth anomaly, proposed by Lakonishok et al. (1994; hereafter, LSV), suggests that the future return patterns of "value" and "growth" firms are due to investors over-extrapolating realized accounting measures of growth. LSV suggest that investors over-react to realized earnings and sales, bidding the price of firms with high past accounting growth too high. In subsequent periods when investor expectations of sales and earnings are not met, the book-to-market ratios of "growth" firms increase as investors reassess their earnings expectations and bid prices down.

In a recent study, Daniel and Titman (2006) provide empirical evidence countering the claims of LSV. DT show that past accounting growth, when measured on a per-share basis, is unrelated to future return patterns. Further, DT show that the future return pattern LSV associate with five-year past accounting growth is actually due to information that is unrelated (i.e., orthogonal) to past accounting returns. This information, tagged intangible returns by DT, is strongly associated with future returns over a five-year period and can be viewed as a proxy for changes in investor expectations of future earnings and/or changes in firm risk.

The empirical evidence of DT raises the question as to whether inferences between other accounting measures associated with firm growth and future returns are also affected by intangible returns and/or equity issuance activity. Of particular interest to the accounting researchers over the past 15 years is the negative relation between future period returns and current period accruals, commonly known as the accrual anomaly. The primary intuition used to explain the accrual anomaly is similar in many respects to that proposed by LSV. Specifically, firms with high (low) current accruals tend to experience lower (higher) future earnings and, because investors do not understand this relation, lower (higher) returns are realized in future periods. However, while the intuition used to explain the future return pattern of the value/growth anomaly is similar to that used to explain the accrual anomaly, the empirical patterns are distinct. Whereas the value/growth anomaly is relatively long in horizon (up to five years), the accrual anomaly is relatively short, with current period accruals only explaining future return variation for about one year. In addition, prior accrual studies show that simple controls for value/growth characteristics (B/M) or risk factors do not explain the accrual anomaly. Accordingly, with the exception of Desai et al. (2004), most studies have viewed the value/growth anomaly and the accrual anomaly as distinct.

Examining the evolution of B/M offers an alternative and possible improvement to the two-period research designs used in prior accrual studies, while examining possible common links between the accrual anomaly and value/growth anomaly. By construction, B/M represents the intersection of four variables: (1) retained earnings, (2) equity investment, (3) expected future earnings, and (4) expected returns. Accordingly, changes in B/M capture changes in at least one of these elements. While the accounting literature has focused on understanding the relation between accrual-based changes in retained earnings and future returns, the finance literature has focused on understanding the relations between changes in the other three elements of B/M and future returns. For example, equity issuance and repurchase activity affect the equity base of the firm, and prior studies have shown that these activities are associated with future returns for up to five years (Loughran and Ritter 1995; Ikenberry et al. 1995). Further, the primary argument in the value/

growth literature is whether the positive association between future returns and price-scaled accounting variables is due to reversals of irrational future earnings expectations formed in prior periods (LSV; Barberis et al. 1998) or rational shifts in risk (Fama and French 1992, 1995).

Prior accrual studies have not necessarily ignored the effects that changes in the other elements of B/M have on inferences between future returns and accruals. Most studies use B/M as a control variable for characteristics associated with firm growth and expected returns. However, it is possible that B/M is too noisy of a variable to control for effects on future returns due to changes in equity issuance, changes in expected earnings, and shifts in risk (Fama and French 2008). To illustrate, consider a firm that experiences a positive shock to expected cash flows in period one, the period prior to the recognition of abnormally large accruals. Due to the discrete, conservative nature of accounting, the shock to future cash flows will tend to be reflected in market value earlier than book value, especially for positive shocks, and B/M will be lower at the end of period one.[5] In the subsequent period, period two, managers capitalize on these positive NPV projects and expand productive capacity. To the extent that managers do not fund these projects with existing cash or debt facilities, managers may issue equity. With this decision to expand productive capacity, positive accruals are recognized. Since accruals are just one element of the four comprising B/M, properly controlling for changes in the other three elements is necessary to distill the incremental future return variation that is explained by accruals from that explained by changes in the other B/M elements occurring in periods contemporaneous with, and prior to, the recognition of accruals.

A decomposition of current B/M into different components offers a structural way to define distinct variables that capture changes in each of these elements. In log form, the book-to-market ratio at time t is equal to the book-to-market ratio at time $t-\tau$ plus log change in book value, minus log change in market value.

$$\begin{aligned}\log(B_t/M_t) = bm_t &= bm_{t-\tau} + \log(B_t/B_{t-\tau}) - \log(M_t/M_{t-\tau}) \\ &= bm_{t-\tau} + \Delta\log(B)_{t-\tau,t} - \Delta\log(M)_{t-\tau,t}.\end{aligned} \tag{1}$$

Log change variables have convenient additive properties that allow a cumulative multi-period change variable to be decomposed into multiple single-period change variables. Using this property, bm_t can be decomposed into the following five year decomposition:

$$bm_t = bm_{t-5} + \log\left(\frac{B_{t-1}}{B_{t-5}}\right) + \log\left(\frac{B_t}{B_{t-1}}\right) - \log\left(\frac{M_t}{M_{t-5}}\right); \tag{2}$$

$$bm_t = bm_{t-5} + \log\left(\frac{B_{t-1}}{B_{t-5}}\right) + \log\left(\frac{NOA_t}{NOA_{t-1}}\right) + \log\left(\frac{\frac{B_t}{NOA_t}}{\frac{B_{t-1}}{NOA_{t-1}}}\right) - \log\left(\frac{M_t}{M_{t-5}}\right)$$

$$\underbrace{\underbrace{\qquad\qquad}_{\text{accruals}\ TACC}\ \underbrace{\qquad\qquad}_{\text{non accrual growth}\ N\_TACC}}_{\text{total log change in } B_t} \tag{3}$$

[5] This tendency is documented by Kothari et al. (2006, Table 2). Their evidence shows that firms classified as high-accrual firms tend to experience increasingly larger abnormal returns in the years leading up to classification as a high-accrual firm. Complementing this evidence, they show that firms classified as low-accrual firms tend to experience increasingly negative abnormal returns in the years leading up to classification as a low-accrual firm, although the absolute magnitude of the returns is smaller relative to the high-accrual firms.

The above decomposition (3) shows that bm_t is equal to bm_{t-5} plus several sub-period, accounting-based log changes in value, minus a cumulative market-based log change. Note that logarithmic properties allow me to decompose log change in book value between $t-1$ and t into the accrual and non-accrual component of log book growth. The definition of *TACC* is consistent to accruals as described in Feltham and Ohlson (1995) and with total accrual variables used in Richardson et al. (2005, 2006), Dechow et al. (2008), and other recent studies examining total accruals.[6,7] *N_TACC* is mathematically equivalent to $\Delta\log(B)_{t-1,t} - \Delta\log(NOA)_{t-1,t}$. Assuming the firm has no equity issuances/repurchases between $t-1$ and t, change in book value equals net income, and *N_TACC* can be interpreted in a manner similar to free cash flow, the cash component of earnings that Dechow et al. (2008) suggest investors also misprice. For robustness purposes, alternative accrual and non-accrual definitions are examined in later sections.

Log Returns versus Log Changes

The log-linear decomposition of the book-to-market ratio is useful because it gives simple variables that can proxy for the dynamics of B/M evolution discussed above. Consistent with DT, I interpret changes in accounting values as representing tangible information (e.g., *TACC*, *N_TACC*). To proxy for intangible information, or information not captured by summary accounting measures, I need to construct a variable that represents information used by investors that is unrelated to accounting information. To do this, I regress realized log market returns over a given period of time on tangible return components of B/M over that same time period. Assuming log market return is a good proxy for total value-relevant information, I interpret the fitted value from this regression as tangible information and the residual as intangible information.

Before running these regressions, log change values defined in Equation (3) above need to be transformed into log return values. Assuming there are no equity issues, dividends, or other corporate actions affecting shareholder's equity, log change in market value between $t-\tau$ and t is equal to the log return an investor earns between $t-\tau$ and t (i.e., $\Delta\log(M)_{t-\tau,t} = ret^{m}{}_{t-\tau,t}$). However, as the time between $t-\tau$ and t increases, $\Delta\log(M)_{t-\tau,t}$ will generally not equal $ret^{m}{}_{t-\tau,t}$. To convert log change in market value into a log return metric, I introduce the variable $iss_{(t-\tau,t)}$. This composite equity issuance variable, $iss_{(t-\tau,t)}$, is defined consistent with the composite equity issuance variable $\iota_{(t-\tau,t)}$ of DT and represents the difference between log market return and total log change in market value. More specifically:

$$iss_{(t-\tau,t)} = \Delta\log(M)_{t-\tau,t} - ret^{m}{}_{t-\tau,t}. \tag{4}$$

The $iss_{t-\tau,t}$ variable can be used to convert $\Delta\log(B)_{t-\tau,t}$ into a log return measure. For example, just as $ret^{m}{}_{t-\tau,t} = \log(\Delta M)_{t-\tau,t} - iss_{t-\tau,t}$, a log book return variable can be constructed as $ret^{b}{}_{t-\tau,t} = \log(\Delta B)_{t-\tau,t} - iss_{t-\tau,t}$. The variable $ret^{b}{}_{t-\tau,t}$ gives the log book value at t of a \$1 investment in book equity made at $t-\tau$. Since $ret^{b}{}_{t-\tau,t}$ is equal to $\Delta\log(B)_{t-\tau,t}$ minus $iss_{t-\tau,t}$, the one-year log book return between $t-1$ and t can be defined as follows:

$$ret^{b}{}_{t-1,t} = \Delta\log(B)_{t-1,t} - iss_{t-1,t} = TACC + N\_TACC - iss_{t-1,t}. \tag{5}$$

---

[6] Richardson et al. (2005, 2006) argue that change in net operating assets offers a more complete measure of accruals. For robustness, this study also considers the more common working capital accruals in Table 4.

[7] Prior studies have generally defined accruals as the linear change in accrual accounts, scaled by total assets. This study defines accruals as the log change in the accrual account (e.g., $\Delta\log(NOA)_{t-1,t}$, $\Delta\log(Working\ Capital)_{t-1,t}$. Conceptually, both types of accrual definitions capture the dynamics associated with one-year growth in net operating assets or net working capital. However, the logarithmic definition can be interpreted as the log return earned by equity holders (i.e., the accrual component of earnings) once equity issuance is controlled. For robustness purposes, the linear variable is directly examined in Table 5.

As Equation (5) reflects, the log book return between $t-1$ and t can be interpreted as the sum of three log return components: total accruals (*TACC*), non-accrual growth (*N_TACC*), and equity issuance.[8] When these three components are used as explanatory variables of future monthly returns, the coefficient on *TACC* will represent the expected average monthly return in period $t+1$ for an investor who realized \$1 of log accrual return between period $t-1$ and t, holding constant log book return attributable to *N_TACC* and $iss_{t-1,t}$. Alternatively, the coefficient on the $iss_{t-1,t}$ variable will capture the expected average monthly return in period $t+1$ for an investor who invested \$1 in the new equity issued at $t-1$.[9]

Intangible Value Change

Once log change in book value ($\Delta\log(B)_{t-\tau,t}$) is transformed into a log book return ($ret^{b}_{t-\tau,t}$), I construct a proxy for intangible information by employing a simple process of estimating cross-sectional annual regressions, defining the residuals from each regression as intangible returns. More specifically, to compute the change in value due to intangible information over the one year current period ($t-1$, t) and the four year prior period ($t-5$, $t-1$), I regress the realized log market returns inclusive of dividends, $ret^{m}_{i,(t-\tau,t)}$, on $bm_{i,(t-\tau)}$ and the components of the $ret^{b}_{i,(t-\tau,t)}$ over the same period of time ($t-\tau$, t).

$$ret^{m}_{i,(t-1,t)} = \gamma_0 + \gamma_1 bm_{i,(t-1)} + \gamma_2 TACC_i + \gamma_3 N\_TACC_i + \gamma_4 iss_{i,(t-1,t)} + u_{i,t}; \quad (6a)$$

$$ret^{m}_{i,(t-5,t-1)} = \gamma_0 + \gamma_1 bm_{i,(t-5)} + \gamma_2 \Delta\log(B)_{i,(t-5,t-1)} + \gamma_3 iss_{i,(t-5,t-1)} + u_{i,t-1}. \quad (6b)$$

Consistent with DT, I define the return due to tangible information as the fitted component and intangible information as the residual from the above regressions:[10]

$u_{i,t} = CPIR$ = current period intangible return; and
$u_{i,t-1} = PPIR$ = prior period intangible return.

By construction, total log return is equal to the sum of tangible returns *plus* intangible returns.[11]

III. RESULTS

Data Definitions and Construction

My study examines the relation between future returns and accruals in firms with fiscal year-ends between 1/31/67 and 12/31/06. To be included in the sample, firms must appear on the CRSP/Compustat merged database and have positive book value of equity at fiscal year-end for

---

[8] There are two alternative methods for transforming total log change in book value into $ret^{b}_{t-\tau,t}$. First, $ret^{b}_{t-\tau,t}$ can be computed using a share adjustment factor as noted in the CRSP/Compustat data manual. Second, $iss_{t-\tau,t}$ can be computed as a "plug:" $iss_{t-\tau,t} = bm_t - bm_{t-\tau} - \Delta\log(B)_{t-\tau} + ret^{m}_{t-\tau}$.

[9] An important subtlety to note is that *TACC* and *N_TACC* can be interpreted as accrual and non-accrual components of earnings in multivariate settings that include $iss_{t-1,t}$. Variation in $\Delta\log(B)_{t-1,t}$ is explained by earnings for that period once equity issuance and repurchase activity is controlled. Hence, in multivariate settings that include $iss_{t-1,t}$, *TACC* can be viewed as the total accrual component of earnings.

[10] Fitted values for 6a and 6b explain 17 percent and 40 percent of the cross-sectional variation of $ret^{m}_{t-1,t}$ and $ret^{m}_{t-5,t-1}$, respectively.

[11] Another interpretation of $iss_{t-\tau,t}$ is that this variable captures elements of intangible information not captured by *PPIR* or *CPIR*. By construction, the intangible returns defined above (*CPIR* and *PPIR*) only capture the intangible information that investors use to price securities. Accordingly, it is reasonable to assume that a manager's intangible information is more precise and extensive than the empirical proxies defined above. Prior literature (Hovakimian et al. (2001)) suggest that firms are more likely to issue (repurchase) equity following periods when stock prices increase (decrease) more dramatically than earnings. Accordingly, $iss_{t-\tau,t}$ can be interpreted as an additional (and possibly independent) measure of intangible information.

years $t-5$, $t-1$, and t. However, firms with negative book values in intermediate years are not excluded. In addition, firms must have positive net operating asset balances at fiscal year-end of years $t-1$, and t. Additional accounting measures used later in robustness tests affect the sample sizes used in those robustness tests accordingly, although not significantly. I exclude financial firms (SIC 6000–6999), utilities (SIC 4900–4999), firms with stock prices on CRSP of less than $5 per share as of the last day of trading in June of year $t+1$, and firms without a complete series of (60) monthly returns between fiscal year-end $t-5$ and fiscal year-end t. The future return series begins on July 1, $t+1$ and delisting returns are used, when necessary, to compute annual portfolio returns.[12,13] I also require that the firm have a valid price and shares outstanding as of the last day of trading in fiscal year $t-5$, $t-1$, t and the last day of trading in June of year $t+1$. From this final set, I annually trim (0.5, 99.5) the tangible return components (*TACC*, *WCACC*, *N_TACC*, and *N_WCACC* variables). Consistent with prior studies, scaled linear change variables used in later robustness tests are annually winsorized (0.5, 99.5).[14] Returns are from CRSP. All other variables are from the CRSP/Compustat merged database.

A firm's log book-to-market ratio at time t (bm_t) is computed as the log book value of the firm at fiscal year-end of year t minus the log market value of the firm as of the last day of trading in fiscal year t as reported in CRSP/Compustat merged database (data25 * data199). All log change variables are defined with two subscripts: the first subscript notes the beginning of the change period and the second subscript notes the end of the change period. For example, $\Delta\log(B)_{t-\tau,t}$ is computed as log book value as of the end of fiscal year t minus the log book value as of the end of fiscal year $t-\tau$. Similarly, log return measures (e.g., $ret^{b}_{t-\tau,t}$) are italicized and defined with two subscripts, with the first subscript noting the beginning of the return period and the second subscript noting the end.

Data Summary

Table 1 reports summary statistics. These statistics show that accruals appear to be a relatively symmetric variable, although there is some right skewness. The skewness is consistent with the average firm growing its balance sheet over time. Given the data requirements of a trailing five-year return series, my sample is skewed toward larger, more mature firms. However, the standard deviation of the variables suggests considerable variation, despite their relative homogeneous size characteristic.

Table 2 reports pairwise correlation coefficients. Overall, the univariate correlations support my hypothesis that *PPIR* affects inferences of the relation between future returns and current period accruals. First, consistent with the prior accrual literature, each accrual measure is negatively correlated with bm_t. Second, *TACC* is more strongly correlated with *PPIR* than working capital accruals (*WCACC*). Further the relation between accrual variables and *PPIR* is stronger than the relation between the respective accrual variables and bm_t, suggesting that bm_t may be an inadequate control for prior period intangible returns. Finally, future returns (*Ret*) are negatively related to each accrual measure as well as *PPIR* and both issuance variables. Given the strong positive association among the accrual measures, *PPIR*, and the issuance variables, the return

[12] Size-adjusted returns are computed as the difference between the raw return as reported on CRSP and the size-matched return reported on the CRSP market capitalization decile return file (crsp.ermport1).

[13] Consistent with DT, I begin the future return series in July of year $t+1$ for all firms, regardless of the month in year t that their fiscal year-ends. Results are qualitatively identical when only firms with December fiscal year-ends are included. If a monthly return is missing during a return period because there is no price on CRSP (a code of ".B" in the WRDS SAS data set), the return is set to zero.

[14] Prior accrual studies have primarily defined accruals as a linear difference, scaled by average total assets (e.g., $NOA_t - NOA_{t-1}$, divided by average total assets). To be consistent with these studies, I winsorize linear difference variables, although results are not overly sensitive to this treatment.

TABLE 1

Summary Statistics

	BM	*ShrsOut*	*Price*	*ME*	*TACC*	*WCACC*	$iss_{t-1,t}$	*Ret*
For All 61,241 Firms								
Mean	0.795	55.35	25.76	2,125.7	0.10	0.08	−0.0001	0.16
Std. Dev.	0.596	250.80	27.84	11,221.3	0.33	0.40	0.1180	0.57
Q25	0.402	4.05	11.38	56.47	−0.05	−0.08	−0.0379	−0.14
Q50	0.644	11.15	19.75	216.41	0.07	0.08	−0.0106	0.08
Q75	1.011	33.98	33.00	905.5	0.22	0.24	0.0122	0.35
For 1968: 633 Firms								
Mean	0.604	12.03	42.75	691.0	0.15	0.16	−0.0010	−0.01
Std. Dev.	0.391	30.39	29.45	2,438.4	0.21	0.33	0.0871	0.33
Min.	0.043	0.11	5.25	3.71	−0.50	−1.00	−0.5583	−0.62
Max	3.919	541.19	353.75	39,700.7	1.35	2.02	0.9316	4.05
For 1977: 1,179 Firms								
Mean	1.142	14.93	22.88	538.6	0.02	0.04	−0.0233	0.21
Std. Dev.	0.637	33.78	17.84	2,098.6	0.27	0.32	0.0671	0.35
Min.	0.164	0.31	5.00	1.57	−1.33	−1.83	−0.6398	−1.00
Max	5.838	615.53	264.0	39,213.2	1.47	1.72	0.5913	1.83
For 1978: 1,457 Firms								
Mean	1.168	13.35	22.56	443.4	0.12	0.12	−0.0238	0.21
Std. Dev.	0.665	32.13	16.12	1,791.6	0.26	0.32	0.0724	0.39
Min.	0.129	0.25	5.00	3.00	−1.44	−1.91	−0.9746	−1.00
Max	8.060	652.88	257.25	39,091.4	1.35	2.06	0.7549	4.28
For 2000: 1,811 Firms								
Mean	0.654	113.82	25.80	5,130.2	0.14	0.10	0.0043	0.14
Std. Dev.	0.491	432.09	27.79	24,959.4	0.38	0.43	0.1438	0.63
Min.	0.006	0.01	5.00	0.07	−1.35	−2.18	−1.96	−1.00
Max	3.710	9,893.43	520.00	524,351.6	2.40	2.19	1.51	6.52

This table reports the mean, standard deviation, lower quartile, median, upper quartile, minimum, and maximum for the primary sample of firm-year observations for firms with fiscal years ending t (1/31/67 and 12/31/06). *BM* is the book-to-market ratio computed as book value of common equity as of fiscal year-end scaled by market value of equity as of fiscal year-end (data25 * data199). *ShrsOut* is equal to the number of shares outstanding (in millions). *Price* is equal to the price per share. *ME* is equal to market value of equity. *TACC* is equal to total accruals computed as $log(NOA_t / NOA_{t-1})$. *WCACC* is equal to working capital accruals computed as $\log(WC_t / WC_{t-1})$. WC_t is computed as current assets (data4) minus cash and short-term investments (data1) minus current liabilities (data5) plus debt in current liabilities (data34) at fiscal year-end t. Specifically, working capital equals $(\text{data4} - \text{data1})_t - (\text{data5} - \text{data34})_t$. $iss_{t-1,t}$ is a composite equity issuance variable computed as the difference between log change in the market value of equity between fiscal year-ends $t-1$ and t and the log market return (inclusive of dividends) between fiscal year-ends $t-1$ and t. Specifically, $iss_{t-1,t} = log(ME_t / ME_{t-1}) - Ret^m_{t-1,t}$. *Ret* equals the buy-and-hold return, inclusive of distributions, between July 1 of year $t+1$ to June 30 of year $t+2$. See the Appendix for the Compustat data items that comprise book value of common equity and *NOA*. *ME*, *ShrsOut*, and *Price* are per CRSP and as of last day of trading in June, $t+1$. *BM*, *TACC*, *WCACC*, and $iss_{t-1,t}$ are of fiscal year-end t and are per the merged CRSP/Compustat database.

reversal that characterizes the accrual anomaly may be due to investor reaction to information not measured by accruals, but correlated with the recognition of accruals.

Regression Results

The null hypothesis of this study is that future period returns are unrelated to realized accounting values such as accruals, consistent with the tenets of market efficiency and risk-neutral

TABLE 2

Full Sample Correlation Statistics

	bm_t	*PPIR*	$iss_{t-5,t-1}$	$iss_{t-1,t}$	*N_TACC*	*TACC*	*WCACC*	*Ret*
bm_t	**1.000**	**−0.587**	**−0.146**	**−0.155**	−0.003	**−0.123**	**−0.106**	**0.086**
PPIR	**−0.578**	**1.000**	0	**0.052**	**0.024**	**0.188**	**0.151**	**−0.027**
$iss_{t-5,t-1}$	**−0.209**	**0.031**	**1.000**	**0.208**	−0.002	**0.089**	**0.071**	**−0.049**
$iss_{t-1,t}$	**−0.284**	**0.107**	**0.470**	**1.000**	**0.149**	**0.206**	**0.155**	**−0.036**
N_TACC	**−0.023**	−0.008	0.001	**0.140**	**1.000**	**−0.738**	**−0.291**	**0.028**
TACC	**−0.160**	**0.229**	**0.106**	**0.159**	**−0.703**	**1.000**	**0.499**	**−0.034**
WCACC	**−0.149**	**0.189**	**0.099**	**0.163**	**−0.403**	**0.641**	**1.000**	**−0.010**
Ret	**0.102**	**−0.023**	**−0.100**	**−0.096**	**0.040**	**−0.041**	**−0.023**	**1.000**

Correlations significant at less than 0.01 level are bolded.

This table reports full sample univariate correlations for log-linear return components of the book-to-market ratio with Pearson (Spearman) coefficients reported above (below) the diagonal. Sample includes firms with fiscal year-ends between 1/31/67 and 12/31/06. bm_t equals the log book-to-market ratio as of fiscal year-end t with book value of common equity and market value of equity as of fiscal year-end t. *PPIR* is equal to prior period intangible (*log*) return, defined as the residual from annual regressions of log market return, $Ret^{m}_{t-5,t-1}$, on $iss_{t-5,t-1}$, $log(B_{t-1}/B_{t-5})$, and bm_{t-5}. $iss_{t-\tau,t}$ is a composite equity issuance variable computed as the difference between log change in the market value of equity between fiscal year-ends $t-\tau$ and t and the log market return (inclusive of dividends) between fiscal year-ends $t-\tau$ and t. *N_TACC* is equal to $log((B_t/NOA_t)/(B_{t-1}/NOA_{t-1}))$. *TACC* is equal to total accruals computed as $log(NOA_t/NOA_{t-1})$. *WCACC* is equal to working capital accruals computed as $log(WCACC_t/WCACC_{t-1})$. *Ret* equals the buy-and-hold return, inclusive of distributions, between July 1 of year t+1 to June 30 of year t+2. See the Appendix for the Compustat data items that comprise book value of common equity and *NOA*.

investors properly incorporating the pricing implications of accruals into price. However, if future period returns are significantly related to current period accruals, and this relation is not explained by risk or other correlated corporate activities such as equity issuance, then the null hypothesis can be rejected. To examine whether accruals are negatively associated with future returns after prior period intangible returns are controlled, I perform incremental regressions that reconcile my results to prior accrual literature and link the accrual anomaly to the value/growth anomaly. I begin by establishing the negative relation between current period accruals and future returns. I then introduce other tangible and intangible components of total market return over the current period. These variables complete the one-year decomposition of the bm_t and largely confirm inferences from prior accrual literature. Finally, I complete the five-year decomposition of bm_t and provide evidence that future return reversal is not associated with current period accruals, but rather the intangible returns that precede accrual recognition. Further, the five-year decomposition of bm_t provides evidence that the accrual anomaly is a function of the value/growth anomaly.

Table 3 reflects the results of Fama-MacBeth regressions of raw monthly returns on components of log book-to-market.[15] Regression 1 of Table 3 shows that, in a univariate specification, total accruals (*TACC*) are strongly negatively associated with future returns, consistent with prior studies. Regression 2 shows that the negative relation holds after controlling for non-accrual growth (*N_TACC*). Regression 3 shows that, even after controlling for equity issuance and non-accrual growth, there is still a strong negative association between *TACC* and future returns.

Regression 4 includes current period intangible returns (*CPIR*) as an additional explanatory

[15] Inferences from empirical analysis are qualitatively the same if size-adjusted returns are examined.

TABLE 3

Regressions of Future Monthly Returns on Log Return Components of Current Book-to-Market

Model	Int.	*TACC*	*N_TACC*	$iss_{t-1,t}$	*CPIR*	*PPIR*	$iss_{t-5,t-1}$	*PPBG*	bm_{t-5}
1	1.245***	−0.598***							
	(5.20)	(−7.52)							
2	1.263***	−0.765***	−0.217						
	(5.34)	(−4.26)	(−1.34)						
3	1.216***	−0.398**	0.108	−1.475***					
	(5.00)	(−2.24)	(0.65)	(−4.75)					
4	1.217***	−0.399**	0.109	−1.478***	−0.317***				
	(5.00)	(−2.25)	(0.65)	(−4.75)	(−2.61)				
5	1.220***	−0.386**	0.106	−1.235***	−0.338***		−0.390***		
	(5.00)	(−2.22)	(0.65)	(−4.65)	(−2.84)		(−3.04)		
6	1.223***	−0.374**	0.111	−1.259***	−0.330***		−0.372***	−0.009	
	(4.98)	(−2.25)	(0.70)	(−4.76)	(−2.77)		(−2.72)	(−0.15)	
7	1.214***	−0.350**	0.135	−1.253***	−0.318***		−0.364***	0.035	0.080*
	(4.97)	(−2.12)	(0.85)	(−4.73)	(−2.67)		(−2.67)	(0.59)	(1.72)
8	1.190***	0.023	0.402***	−1.34 ***	−0.337***	−0.284***	−0.376***	0.023	0.088*
	(4.89)	(0.14)	(2.64)	(−5.02)	(−2.84)	(−4.06)	(−2.73)	(0.39)	(1.88)
9	1.187***	−0.062	0.349**	−1.566***	−0.336***	−0.268***			
	(4.91)	(−0.38)	(2.24)	(−5.04)	(−2.77)	(−3.96)			

*, **, *** Refers to significance at the 10 percent, 5 percent, and 1 percent levels, two-tailed, respectively.

This table reports the results of a set of Fama-MacBeth regressions of monthly returns on components of a log-linear return decomposition of the book-to-market ratio at fiscal year-end t. The monthly return series is between July of year t+1 and June of year t+2 and t is between 1/31/67 and 12/31/06. All coefficients are multiplied by 100. t-statistics are reported in parentheses. bm_{t-5} is the log book-to-market ratio at fiscal year-end t–5. *TACC* is equal to total accruals computed as $log(NOA_t / NOA_{t-1})$. *N_TACC* is equal to $log((B_t / NOA_t) / (B_{t-1} / NOA_{t-1}))$. $iss_{t-\tau,t}$ is a composite equity issuance variable computed as the difference between log change in the market value of equity between fiscal year-ends $t-\tau$ and t and the log market return (inclusive of dividends) between fiscal year-ends $t-\tau$ and t. Specifically, $iss_{t-1,t} = log(ME_t / ME_{t-1}) - Ret^{m}_{t-1,t}$. *CPIR* is the current period intangible return, defined as the residual from annual regressions of log market return, $Ret^{m}_{t-1,t}$, on bm_{t-1}, *TACC*, *N_TACC*, and $iss_{t-1,t}$. *PPIR* is the prior period intangible return, defined as the residual from annual regressions of log market return, $Ret^{m}_{t-5,t-1}$, on bm_{t-5}, $log(B_{t-1} / B_{t-5})$, and $iss_{t-5,t-1}$. *PPBG* is the prior period log book growth computed as $log(B)_{t-1} - log(B)_{t-5}$. See Appendix for the Compustat data items that comprise book value of common equity and *NOA*.

variable.[16] The significant negative relation between future period returns and *CPIR* suggests that, in expectation, current period returns are not a strict function of current period changes in book value. This finding is important because many accrual studies that adopt the Mishkin model implicitly assume that, in expectation, only earnings and earnings components explain future returns. As Kraft et al. (2007) show, to the extent this assumption is incorrect and the omitted explanatory variable is mispriced, incorrect inferences can result regarding the relation between future period returns and current period accruals.[17]

Subsequent specifications (5–7) introduce tangible return components from the four-year period prior to accrual recognition. Specifically, regressions 5–7 introduce prior period equity issuance ($iss_{t-5,t-1}$), prior period log book growth (*PPBG*), and long-horizon expectations (bm_{t-5}). Inferences on these additional explanatory variables are consistent with those of DT. More importantly, even with the current and prior period log returns controlled, there is still an economically and statistically strong negative relation between future returns and current period accruals. However, once *PPIR* is introduced as an explanatory variable in regression 8, completing the log-linear return decomposition of bm_t, accruals are no longer associated with future returns at conventional levels of significance.

Building off regression 4, which shows that *TACC* are associated with future returns after current period tangible and intangible returns are controlled, regression 9 adds *PPIR* as a control variable. Regression 9 shows that, after *PPIR* is controlled, *TACC* is unrelated with future returns. This finding implies that *PPIR* is a powerful control variable that can distort inferences of current period accounting growth metrics and future returns when omitted.

IV. ALTERNATIVE VARIABLE SPECIFICATIONS AND ROBUSTNESS TESTS

Alternative Accrual Definitions

Table 4 examines other measures of accruals to determine whether the exclusion of prior period intangible returns in prior studies explains the relation between accruals and future returns. The first set of regressions in Table 4, Panel A examines the relation between the components of total accruals suggested by Richardson et al. (2006) and future returns. Richardson et al. (2006) linearly decompose total accruals into three parts: change in sales ($\Delta Sales$), change in asset turnover (ΔATO), and an interaction between $\Delta Sales$ and ΔATO. Using a log-linear decomposition avoids the complicating interaction term, yielding log total accruals (*TACC*) as the log sum of $\Delta\log(Sales)_{t-1,t}$ and $\Delta\log(ATO)_{t-1,t}$.

Consistent with the intuition of Richardson et al. (2006), Table 4, Panel A shows a strong statistical and economic relation between both $\Delta\log(Sales)_{t-1,t}$ and $\Delta\log(ATO)_{t-1,t}$ and future returns.[18] Subsequent controls for current and prior period log return components in regressions 2 and 3 show that $\Delta\log(ATO)_{t-1,t}$ is still significantly associated with future returns, although $\Delta\log(Sales)_{t-1,t}$ becomes insignificant with the introduction of $iss_{t-1,t}$.[19] However, once *PPIR* is added as an explanatory variable in regression 4, completing the log-linear decomposition of bm_t, neither accrual component is associated with future returns.

---

[16] Since some firms used to compute *CPIR* do not have a full series of 12 future monthly returns, the coefficients on *TACC*, *N_TACC*, and $iss_{t-1,t}$ change slightly with the introduction of *CPIR*.

[17] Alternatively, to the extent that the omitted variable is correctly priced, but the expected return model is incorrectly specified, incorrect inferences can also result.

[18] Richardson et al. (2006) do not directly examine the relation between future returns and accruals. However, the negative relation between future returns and accrual components reflected in Table 4, Panel A is consistent with the persistence coefficients reported in their study.

[19] The lack of a significant relation between the "growth" component of accruals, $\Delta\log(Sales)_{t-1,t}$, and future returns is consistent with the results of Daniel and Titman (2006).

TABLE 4

Regressions of Future Monthly Returns on Log Return Components of Current Book-to Market: Alternative Accrual Specifications

Panel A: Total Accrual Components

Model	Int.	$\Delta log(Sales)_{t-1,t}$	$\Delta log(ATO)_{t-1,t}$	*N_TACC*	$iss_{t-1,t}$	$CPIR^{d}$	*PPIR*	$iss_{t-5,t-1}$	*PPBG*	bm_t
1	1.263***	−0.729***	−0.717***	−0.185						
	(5.35)	(−3.34)	(−4.34)	(−1.24)						
2	1.214***	−0.346	−0.386**	0.123	−1.462***	−0.334***				
	(4.99)	(−1.58)	(−2.32)	(0.79)	(−4.62)	(−2.75)				
3	1.215***	−0.288	−0.354**	0.133	−1.222***	−0.339***		−0.361***	0.026	0.077*
	(4.97)	(−1.41)	(−2.28)	(0.88)	(−4.58)	(−2.84)		(−2.63)	(0.44)	(1.66)
4	1.188***	0.115	0.005	0.391***	−1.323***	−0.358***	−0.290***	−0.373***	0.013	0.086*
	(4.89)	(0.58)	(0.03)	(2.66)	(−4.92)	(−3.00)	(−4.14)	(−2.71)	(0.22)	(1.82)
5	1.182***	0.016	−0.061	0.356**	−1.564***	−0.353***	−0.273***			
	(4.89)	(0.08)	(−0.39)	(2.41)	(−4.95)	(−2.90)	(−4.05)			

Panel B: Working Capital Accruals

Model	Int.	*WCACC*	*N_WCACC*	$iss_{t-1,t}$	$CPIR^{w}$	*PPIR*	$iss_{t-5,t-1}$	*PPBG*	bm_{t-5}
1	1.238***	−0.547***	−0.300*						
	(5.18)	(−3.01)	(−1.73)						
2	1.188***	−0.143	0.082	−1.709***	−0.298**				
	(4.84)	(−0.78)	(0.46)	(−5.23)	(−2.41)				
3	1.202***	−0.104	0.106	−1.46 ***	−0.305**		−0.319**	−0.028	0.058
	(4.85)	(−0.60)	(0.62)	(−5.18)	(−2.52)		(−2.26)	(−0.46)	(1.20)
4	1.173***	0.226	0.387**	−1.527***	−0.327***	−0.285***	−0.336**	−0.023	0.067
	(4.76)	(1.35)	(2.39)	(−5.40)	(−2.69)	(−3.96)	(−2.34)	(−0.39)	(1.37)
5	1.158***	0.167	0.345**	−1.780***	−0.321***	−0.277***			
	(4.73)	(0.97)	(2.07)	(−5.45)	(−2.60)	(−3.91)			

*, **, *** Refers to significance at the 10 percent, 5 percent, and 1 percent levels, two-tailed, respectively.

This table reports the results of a set of Fama-MacBeth regressions of monthly returns on components of a log-linear return decomposition of the book-to-market ratio at fiscal

(continued on next page)

year-end t. The monthly return series is between July of year t+1 and June of year t+2 and t is between 1/31/67 and 12/31/06. All coefficients are multiplied by 100. t-statistics are reported in parentheses. $\Delta log(Sales)_{t-1,t}$ is log change in sales (data12) for fiscal period $t-1$ to t. $\Delta log(ATO)_{t-1,t}$ is log change in asset turnover for fiscal period $t-1$ to t where ATO_t is defined as $NOA_t / Sales_t$. *N_TACC* is equal total non-accrual growth for the fiscal year $t-1$ to t. *N_WCACC* is equal to working capital non-accrual growth for the fiscal year $t-1$ to t defined as $log(B_t / WC_t) / (B_{t-1} / WC_{t-1})$. $CPIR^d$ is the current period intangible return defined as the residual from annual regressions of log market return, $Ret^m{}_{t-1,t}$, on bm_{t-1}, $\Delta log(Sales)_{t-1,t}$, $\Delta log(ATO)_{t-1,t}$, *N_TACC*, $iss_{t-1,t}$. $CPIR^w$ is the current period intangible return defined as the residual from annual regressions of log market return, $Ret^m{}_{t-1,t}$, on bm_{t-1}, *WCACC*, *N_WCACC*, and $iss_{t-1,t}$. *TACC*, *WCACC*, $iss_{t-\tau,t}$, *PPIR*, *PPBG*, and bm_{t-5} are the same as defined in prior tables. See Appendix for Compustat data items comprising book value of common equity and *NOA*.

Panel B of Table 4 examines working capital accruals. Due to the logarithmic nature of this variable, I omit depreciation from my working capital accrual variable. This omission will lead to a positive working capital accruals on average (see Table 1), but should not affect regression inferences.[20] Regression 1 shows the telltale negative relation between *WCACC* and future returns. However, this relation becomes insignificant once $iss_{t-1,t}$ is introduced into the regression specification. Regressions 3–5 reinforce the validity of the relations noted in prior tables; that is, there is a strong negative relation between equity issuance and intangible returns with future returns, while the accrual component of earnings is unrelated with future returns.

The relation between *PPIR* and current period accruals dovetails nicely into the intuition of prior studies that explain the accrual anomaly as a function of firm growth and investor misunderstanding of the diminishing returns on investment that follow firm growth. However, the evidence in this study suggests a different interpretation of the relation between future returns and current period accruals than that proposed in prior studies. While investors may misunderstand the diminishing returns on investment, this misunderstanding is not caused by the information contained in accruals (or other accounting performance measures). Rather, the empirical evidence of this study suggests that the return reversal that prior studies have attributed to accounting measures of growth appears to be due to growth not captured by the financial accounting system, denoted as intangible returns.

Linear Accrual Variable

Prior literature has defined current period accruals as a linear difference scaled by assets. For example, total accruals have been defined as the one-year change in net operating assets scaled by beginning period assets.

$$TACC^{linear} = \frac{(NOA_t - NOA_{t-1})}{TA_{t-1}}.$$

Simple algebra shows that $TACC^{linear}$ can be expressed as the product of a beginning stock measure and a one-year growth measure:

$$TACC^{linear} = \frac{NOA_{t-1}}{TA_{t-1}} \times \left(\frac{NOA_t}{NOA_{t-1}} - 1 \right). \tag{7}$$

Equation (7) is useful because it shows that $TACC^{linear}$ is similar to the log-linear accrual variable (*TACC*) used in this study.[21] Given that $\frac{NOA_{t-1}}{TA_{t-1}}$ is positive for all observations in this study, $TACC^{linear}$ is a monotonic transformation of one-year growth in *NOA*. Thus, $TACC^{linear}$ should capture the same underlying economic dynamics as *TACC*, albeit it in a different scale. $TACC^{linear}$ captures growth on a linear basis, whereas *TACC* captures growth on a log-linear basis. This difference will cause $TACC^{linear}$ values to be more extreme than *TACC* values for observations with the most extreme *positive NOA* growth. Accordingly, differences in the regression coefficients of $TACC^{linear}$ and *TACC* may be due to differences in their respective econometric properties, not the economic dynamics they capture.

Table 5 examines whether the primary results of Table 3 are sensitive to accrual measurement. Consistent with the intuition used in prior literature to explain the accrual anomaly, regression 1 establishes that the component of $TACC^{linear}$ responsible for the strong negative relation with future returns is the one-year growth in *NOA* component. Regression 2 directly examines the sensitivity of the results from Table 3 to the accrual definition used by regressing future monthly

[20] Working capital accruals that include depreciation tend to have negative values, which are undefined in log form.
[21] Empirically, the Pearson and Spearman correlations between *TACC* and $TACC^{linear}$ are (ρ = 0.85 and 0.98).

TABLE 5

Regressions of Future Monthly Returns on Log-Linear Return Components of B/M: Linear Accrual Variables

Panel A: Regressions of Future Monthly Returns on Linear Accrual Variables

Model	Int.	$log\left(\frac{NOA_{t-1}}{TA_{t-1}}\right)$	$NOA\ Growth_{t-1,t}$	$TACC^{linear}$	TACC_RNK	N_TACC	$iss_{t-1,t}$	CPIR	PPIR
1	1.163***	−0.208**	−0.517***						
	(4.74)	(−2.25)	(−7.85)						
2	1.218***			−0.458**		0.193*	−1.283***	−0.333***	−0.240***
	(5.05)			(−2.50)		(1.86)	(−4.16)	(−2.75)	(−3.50)
3	1.265***				−0.0018*	0.273***	−1.470***	−0.331***	−0.251***
	(5.23)				(−1.82)	(2.82)	(−4.90)	(−2.73)	(−3.71)
4	1.368***				−0.004 ***	0.151	−1.484***	−0.313***	
	(5.58)				(−3.50)	(1.49)	(−4.89)	(−2.58)	

Panel B: Regressions of Future Monthly Returns on Linear Accrual Variable—Top 1% Trimmed

	Int.	$TACC^{linear}$	N_TACC	$iss_{t-1,t}$	CPIR	PPIR
5	1.215***	−0.387*	0.222**	−1.397***	−0.334***	−0.234***
	(5.05)	(−1.92)	(2.08)	(−4.17)	(−2.72)	(−3.41)
6	1.236***	−0.787***	0.083	−1.377***	−0.318***	
	(5.09)	(−3.75)	(0.75)	(−4.09)	(−2.59)	

Panel C: Regressions of Future Monthly Returns on Linear Accrual Variable—Top 2% Trimmed

	Int.	$TACC^{linear}$	N_TACC	$iss_{t-1,t}$	CPIR	PPIR
7	1.214***	−0.286	0.249**	−1.320***	−0.333***	−0.241***
	(5.00)	(−1.35)	(2.27)	(−3.62)	(−2.72)	(−3.51)
8	1.235***	−0.699***	0.107	−1.305***	−0.316***	
	(5.07)	(−3.13)	(0.94)	(−3.57)	(−2.57)	

*, **, *** Refers to significance at the 10 percent, 5 percent, and 1 percent levels, two-tailed, respectively.

(continued on next page)

This table reports the results of Fama-MacBeth regressions of future monthly returns on tangible and intangible return components with accrual variables defined in linear form. The monthly return series is between July of year t+1 and June of year t+2 and t is between 1/31/67 and 12/31/06. All coefficients are multiplied by 100. t-statistics are reported in parentheses. *NOA Growth*$_{t-1,t}$ is equal to $(NOA_t - NOA_{t-1}) / NOA_{t-1}$. $TACC^{linear}$ is equal to $(NOA_t - NOA_{t-1})$ divided by total assets (data6) at t−1. *TACC_RNK* is the annual rank (0–99) of $TACC^{linear}$. *N_TACC*, $iss_{t-1,t}$, *CPIR*, and *PPIR* are the same as defined in prior tables. Panel A reports results from full sample. Panel B reports results from LTS specification with top 1 percent of $TACC^{linear}$ observations trimmed per year. Panel C reports results from LTS specification with top 2 percent of $TACC^{linear}$ observations trimmed per year. Consistent with prior accrual literature that adjusts linear accounting variables, $TACC^{linear}$, $log(NOA_{t-1} / TA_{t-1})$, and *NOA Growth*$_{t-1,t}$ are annually winsorized (0.5, 99.5).

returns on current and prior period return components, with $TACC^{linear}$ substituted for *TACC*. The coefficients in regression 2 are largely consistent in economic and statistical significance to those reported in Table 3, except that $TACC^{linear}$ shows a significantly negative association with future returns. Given that regression 1 shows the negative relation between $TACC^{linear}$ and future period returns is due primarily to the one-year growth in *NOA*, which is insignificant when measured logarithmically (*TACC*), the significant coefficient on $TACC^{linear}$ reflected in regression 2 could be due to the way values of $TACC^{linear}$ are affected by extreme growth in *NOA*.

To examine whether the significant negative relation between future returns and $TACC^{linear}$ is due to economic differences or econometric differences between $TACC^{linear}$ and *TACC*, I transform $TACC^{linear}$ into percentile ranks. Specifically, I rank firms annually based on $TACC^{linear}$ (0 to 99) and use these ranks as accrual explanatory variables. This transformation reduces right skewness in the accrual variable, thereby reducing the leverage on the $TACC^{linear}$ regression coefficients caused by observations with extreme realizations of *NOA* growth. This transformation maintains the ordinal ranking of the variable and should not affect inference differences caused by economic differences between *TACC* and $TACC^{linear}$. Regression 3 in Table 5, Panel A reports the results of this transformation with economic and statistical significance consistent with that reported in Table 3. For comparison to prior tables, regression 4 removes *PPIR* as an explanatory variable and shows that, consistent with prior tables, the exclusion of *PPIR* significantly affects inferences between accruals and future returns.

An alternative way to examine the effect that extreme positive growth in *NOA* has on regression inferences between future returns and $TACC^{linear}$ is to use a least trimmed squares (LTS) research design. While OLS regressions have many desirable attributes, they are sensitive to leverage points (e.g., extreme independent variable values). Given that extreme positive growth in *NOA* will cause $TACC^{linear}$ to lie disproportionately on the right-hand side of the x-axis (relative to *TACC*), and thereby impart high leverage on the regression coefficients, I examine the sensitivity of regression inferences to observations with extreme positive $TACC^{linear}$ values. Specifically, I employ a LTS empirical design in the spirit of Knez and Ready (1997) and Kraft et al. (2006) and annually trim firms in the 100th percentile of one-year growth in *NOA*. If $TACC^{linear}$ captures a robust economic relation incremental to that of *TACC*, then its relation with future returns should not be affected by the omission of the most extreme $TACC^{linear}$ observations. However, if the significant negative relation noted in regression 2 of Table 5 is econometric rather than economic and caused by extreme values of $TACC^{linear}$, then removing extreme observations will have a significant effect on inferences.

Panels B and C of Table 5 report the regression results based on the LTS procedure, with Panel B trimming the most extreme $TACC^{linear}$ percentile (top 1 percent) and Panel C trimming the top two extreme $TACC^{linear}$ percentiles (top 2 percent). Results suggest that extreme $TACC^{linear}$ values have a significant effect on regression inferences. Regression 5 shows only a moderately significant relation between future returns when only the top $TACC^{linear}$ percentile is trimmed; regression 7 shows no relation. Equally important are regressions 6 and 8, which show that the exclusion of *PPIR* as a control variable has a significant effect on inferences between future returns and current period accruals.[22]

The evidence in Table 5 suggests that the negative relation between current period accruals and future returns is sensitive to the manner in which accruals are measured. However, this sensitivity is not caused by differences in the economic fundamentals captured by the log-linear accrual variable (*TACC*) versus the linear accrual variable $TACC^{linear}$. Rather, the sensitivity ap-

[22] Inferences from regression specification (2) (Panel A) are qualitatively identical when a LTS procedure is employed that trims monthly observations in the highest squared residual percentile (1 percent) as done in Kraft et al. (2006).

pears to be due to differences in the econometric properties of the two accrual variables. Table 5 reinforces the point raised by Kraft et al. (2006) regarding the influence of extreme observations on return/accrual inferences and highlights the importance of using multiple variable definitions when examining the properties of accruals. Further, Table 5 reinforces that *PPIR* has a significant effect on the relation between current period accruals and future returns.

The Relation between Non-Accrual Growth and Future Returns

The non-accrual book return component, *N_TACC*, is insignificant in regression specifications that exclude *PPIR*, but becomes positive and significant in specifications that include *PPIR*. As discussed above, *N_TACC* effectively captures changes in book value due to free cash flow. However, unlike the traditional free cash flow variables defined in a linear manner (e.g., (Net Income − Accruals)/Avg. Total Assets), *N_TACC* assumes that positive free cash flow reduces a firm's net debt position and negative free cash flow increases a firm's net debt position. In essence, *N_TACC* can be viewed as a lever that amplifies (attenuates) log book returns. For example, if *N_TACC* is positive, then this implies that a firm was able to reduce its net debt position after operating and investing needs were met, thereby amplifying the log book returns realized by equity holders (where net debt equals debt minus cash). On the other hand, if *N_TACC* is negative, it means that some growth in net operating assets was due to an increase in net debt. As such, the log book return is attenuated since part of the book return is funded by an increase in net debt.

To examine the sensitivity of my results to using *N_TACC* as a control variable, Table 6 substitutes two alternative variables that capture the effect on log book return due to debt issuance activities: *ChgDebt* and *ChgNetDebt*.

$$ChgDebt = \log\left(\frac{(B_{t-1} + Debt_t - Debt_{t-1})}{B_{t-1}}\right); \tag{8a}$$

$$ChgNetDebt = \log\left(\frac{(B_{t-1} + NetDebt_t - NetDebt_{t-1})}{B_{t-1}}\right). \tag{8b}$$

The results from Table 6 largely confirm inferences from prior tables. Specifically, the accrual component of log book return is not associated with future returns once current period net debt changes and prior period intangible return components are controlled.[23] However, in specifications that exclude *PPIR*, there is a significant negative relation between the accrual component of earnings and future returns.

The fact that *N_TACC*, $ChgDebt_{t-1,t}$, and $ChgNetDebt_{t-1,t}$ are tangible return components and are significantly associated with future returns does not change the primary contention of this study or that of Daniel and Titman (2006). Richardson et al. (2005) classify the reliability of all balance sheet accruals and suggest that accruals relating to debt instruments are the most reliable and are the least likely to be mispriced. This contention is premised on the fact that managers have very little discretion in the recognition of debt-based accruals, which leads to little measurement error in these variables. Accordingly, if investors are unlikely to misprice the information contained in *N_TACC*, $ChgDebt_{t-1,t}$, and $ChgNetDebt_{t-1,t}$, then a more likely explanation is that these variables have characteristics similar to $iss_{t-\tau,t}$. As discussed in DT, managers are more likely to issue equity in response to positive growth opportunities or in response to a perceived decline in the cost of equity. In either case, the issuance of equity is triggered by intangible returns. In a similar vein, the debt issuance activity of a firm appears also to be triggered by realizations of

[23] For brevity, similar regression analyses examining working capital accruals are not reported, but results are consistent with those reflected in Table 6 for total accruals.

TABLE 6

Regressions of Future Monthly Returns on Log-Linear Return Components of B/M: Alternative Debt Issuance Variables

Panel A: Regressions of Future Monthly Returns, Controlling for Change in Net Debt.

	Int.	*TACC*	*ChgNetDebt*	$iss_{t-1,t}$	*CPIR*	*PPIR*
1	1.195***	−0.153	−0.237***	−1.448***	−0.347***	−0.236***
	(4.97)	(−1.48)	(−3.42)	(−4.87)	(−2.84)	(−3.41)
2	1.209***	−0.327***	−0.185***	−1.510***	−0.332***	
	(5.02)	(−2.94)	(−2.56)	(−5.06)	(−2.72)	

Panel B: Regressions of Future Monthly Returns, Controlling for Change in Debt

	Int.	*TACC*	*ChgDebt*	$iss_{t-1,t}$	*CPIR*	*PPIR*
3	1.216***	−0.223***	−0.291***	−1.294***	−0.325***	−0.221***
	(5.04)	(−2.62)	(−4.37)	(−4.31)	(−2.65)	(−3.18)
4	1.221***	−0.330***	−0.305***	−1.398***	−0.313**	
	(5.05)	(−3.74)	(−4.55)	(−4.59)	(−2.56)	

Panel C: Regressions of Future Monthly Returns, Controlling for Changes in Debt Ranks

	Int.	$TACC^{linear}$ *RNK*	*ChgDebt RNK*	*ChgNetDebt RNK*	$iss_{t-1,t}$ *RNK*	*CPIR RNK*	*PPIR RNK*
5	2.045***	−0.0003		−0.0045***	−0.0053***	−0.0036***	−0.0037***
	(8.89)	(−0.21)		(−3.30)	(−3.08)	(−2.68)	(−2.99)
6	2.024***	−0.0016*	−0.0039***		−0.0048***	−0.0034**	−0.0032**
	(8.56)	(−1.71)	(−4.97)		(−2.89)	(−2.47)	(−2.61)

*, **, *** Refers to significance at the 10 percent, 5 percent, and 1 percent levels, two-tailed, respectively.

This table reports the results of Fama-MacBeth regressions of future monthly returns on tangible and intangible return components. The monthly return series is between July of year t+1 and June of year t+2 and t is between 1/31/67 and 12/31/06. All coefficients are multiplied by 100. t-statistics are reported in parentheses. *TACC* is equal to $log(NOA_t / NOA_{t-1})$. *ChgNetDebt* is equal to $log((B_{t-1} + NFO_t - NFO_{t-1}) / B_{t-1})$, where *NFO* is equal to financial liabilities − financial assets. Financial liabilities equal long-term debt (data9) + short-term debt (data34) + notes payable (data206) + preferred stock (shareholder's equity − book value of equity). Financial assets equal investments and advances (data32) +

(continued on next page)

short-term investments (data193) + cash (data1). *ChgDebt* is equal to $log((B_{t-1} + Debt_t - Debt_{t-1}) / B_{t-1})$, where *Debt* is equal to long-term debt (data9) + short-term debt (data34) + notes payable (data206) + preferred stock (shareholder's equity − book value of equity). $TACC^{linear}$ is equal to $(NOA_t - NOA_{t-1})$ divided by total assets (data6) at $t-1$. *N_TACC*, $iss_{t-1,t}$, *CPIR*, and *PPIR* are the same as defined in prior tables. Ranked variables are equal to the annual rank (0–99) of the respective variable. Consistent with prior accrual literature that adjusts linear accounting variables, *ChgDebt* and *ChgNetDebt* are annually winsorized (0.5, 99.5).

intangible information in current and prior periods. However, note that, even if the relation between the debt issuance variables and future returns is symptomatic of some sort of behavioral pricing bias, the intuition used to explain this pricing bias must be different than that relied upon in prior accrual studies that asserts that greater manager discretion leads to greater measurement error and hence, greater mispricing.

Portfolio Tests

To examine whether my results hold in a nonparametric, portfolio setting, I employ a series of dependent sorts that minimize the variation in the control variables, while maximizing the variation in the variable of interest. The variable I examine is log accrual return ($TACC^{adj}$), and the two control variables are total period intangible return (*TPIR*) and non-accrual growth (*N_TACC*). $TACC^{adj}$ represents growth in net operating assets on a per-share basis (i.e., log accrual return). *TPIR* represents the component of total log return realized between $t-5$ and t that is orthogonal to $TACC^{adj}$, *N_TACC*, and $ret^{b}_{t-5,t-1}$ (where $ret^{b}_{t-5,t-1}$ equals $\log(B_{t-1} / B_{t-5}) - iss_{t-5,t-1}$).

To examine how variation in $TACC^{adj}$ explains future returns, I sort all observations into annual deciles ranked on *TPIR*. Each of these ten annual *TPIR* deciles is then sorted into ten annual sub-deciles based on *N_TACC*. Each of these annual sub-deciles is then sorted into annual quintiles based on $TACC^{adj}$. I then pool the sub-deciles together by year, based on their respective $TACC^{adj}$ rank. The result is five portfolios in each year, with significant variation in accruals and minimal variation across the control variables (e.g., *TPIR* and *N_TACC*). Values reflected in Table 7 represent equal-weighted annual averages of the 40 years in the sample, with t-statistics consistent with Fama and MacBeth (1973). Delisting returns are incorporated into the computation of annual returns as per Beaver et al. (2007), although results are not sensitive to the inclusion of delisting returns.[24]

Panel A of Table 7 shows the average hedge portfolio return (size-adjusted and raw) from implementing an unconditional portfolio strategy of sorting firms each year into accrual deciles based on *TACC* and taking a long (short) position in the low (high) accrual firms. Size-adjusted returns are denoted *sRet* and raw returns are denoted *Ret*. For comparative purposes, I also reflect $TACC^{adj}$ and *WCACC* to show the close similarity between portfolios formed on *TACC* and those formed on $TACC^{adj}$ and *WCACC* (where $TACC^{adj} = TACC - iss_{t-1,t}$).

Panel A of Table 7 suggests that the hedge returns are statistically significant, and the economic significance is comparable to those reflected in prior accrual studies. Panels B, C, and D use the same sample of firms, sorted into quintiles as described above, to determine whether accruals explain future returns after controlling for correlated characteristics.

Panel B of Table 7 shows that once *TPIR* and *N_TACC* are controlled, accruals are not associated with future returns. Panel C in Table 7 examines the abnormal returns that can be earned on *TPIR*, controlling for $TACC^{adj}$ and *N_TACC*. This panel shows that economically and statistically significant abnormal returns can be earned on portfolios sorted on *TPIR*, suggesting that intangible returns capture important cross-sectional differences in either risk or some type of behavioral bias. Panel D in Table 7 examines the abnormal returns that can be earned on *N_TACC* controlling for *TPIR* and $TACC^{adj}$. Similar to Panel B, the average hedge return across the 40 years of the sample is only marginally statistically significant, although economically insignificant.[25]

---

[24] Firms with delisting codes between 500 and 599 (liquidations) receive a delisting return of −1.0. I thank Richard Price for his delisting return code at http://richardp.rice.edu/research/delistings.sas.

[25] Results are qualitatively identical (and quantitatively very similar) if *ChgNetDebt* or *ChgDebt* are substituted for *N_TACC*.

TABLE 7
Portfolio Sorts

Panel A: Portfolios Based on Annual Sorts of *TACC*

Portfolio	Ret_{t+1}	$sRet_{t+1}$	*TACC*	$TACC^{adj}$
1	0.183	0.058	−0.426	−0.405
2	0.156	0.039	−0.123	−0.105
3	0.160	0.039	−0.043	−0.024
4	0.166	0.045	0.009	0.027
5	0.159	0.039	0.052	0.067
6	0.159	0.036	0.097	0.107
7	0.133	0.015	0.150	0.155
8	0.139	0.019	0.218	0.214
9	0.129	0.009	0.333	0.317
10	0.094	−0.021	0.717	0.650
Avg. Hedge Ret.	0.089***	0.079***		
FM t-statistic	4.87	4.62		

Panel B: Portfolios Based on $TACC^{adj}$, Controlling for *N_TACC* and *TPIR*

Portfolio	Ret_{t+1}	$TACC^{adj}$	*N_TACC*	*TPIR*
1	0.139	−0.131	0.054	0.005
2	0.155	0.031	0.026	−0.002
3	0.156	0.104	0.003	−0.003
4	0.148	0.172	−0.016	−0.001
5	0.139	0.308	−0.047	0.001
Avg. Hedge Ret.	−0.000			
FM t-statistic	−0.038			

(continued on next page)

Panel C: Portfolios Based on *TPIR*, controlling for *TACC* and *N_TACC*

Portfolio	Ret_{t+1}	$TACC^{adj}$	*N_TACC*	*TPIR*
1	0.166	0.102	0.011	−0.664
2	0.162	0.099	0.007	−0.290
3	0.155	0.101	0.006	−0.027
4	0.139	0.102	0.003	0.249
5	0.115	0.099	−0.009	0.699
Avg. Hedge Ret.	0.051**			
FM t-statistic	(2.62)			

Panel D: Portfolios Based on *N_TACC*, Controlling for *TACC* and *TPIR*

Portfolio	Ret_{t+1}	$TACC^{adj}$	*N_TACC*	*TPIR*
1	0.133	0.154	−0.243	0.014
2	0.141	0.123	−0.071	−0.004
3	0.150	0.101	0.007	−0.005
4	0.161	0.078	0.075	−0.003
5	0.152	0.049	0.232	0.002
Avg. Hedge Ret.	−0.019*			
FM t-statistic	(−1.75)			

*, **, *** Refers to significance at the 10 percent, 5 percent, and 1 percent levels, two-tailed, respectively.

This table reports the results of hypothetical hedge returns earned on annually rebalanced portfolios. Panel A reports decile portfolios formed annually on 6/30/68 to 6/30/07 for firms with fiscal year-ends between 1/31/67 and 12/31/06. On June 30, t+1, firms are sorted into deciles based on *TACC* from fiscal year-end t. Ret_{t+1} is the 12 month, buy-and-hold raw return beginning in July of year t+1. $sRet_{t+1}$ is the 12 month, buy-and-hold size-adjusted return beginning in July of year t+1. $TACC^{adj}$ are per share total accrual defined as $\log(NOA_t / NOA_{t-1}) - iss_{t-1,t}$. *TPIR* is the total period intangible return, defined as the residual from annual regressions of log market return, $Ret^{m}{}_{t-5,t}$, on bm_{t-5}, *TACC, N_TACC*, $Ret^{b}{}_{t-5,t-1}$, where $Ret^{b}{}_{t-5,t-1}$ is equal to $log(B_{t-1} / B_{t-5}) - iss_{t-5,t-1}$. Other variables are the same as detailed in prior tables. Panels B, C, and D reflect results from hedge portfolios formed on $TACC^{adj}$, *TPIR*, and *N_TACC*, respectively. To maximize the variation in the variable of interest, while minimizing the variation in the two other control variables, I perform a series of dependent sorts. Each calendar year, I sort all firms into annual deciles ranked on the first control variable. Each of these ten annual deciles is then sorted into 10 sub-deciles based on the second control variable. Each of these annual sub-deciles is then sorted into annual quintiles based on variable of interest. I then pool the sub-deciles each year based on the respective variable of interest. The variable of interest in Panel B is $TACC^{adj}$, in Panel C it is *TPIR*, and in Panel D it is *N_TACC*.

V. UNTABULATED ROBUSTNESS TESTS

Inferences are not sensitive to the inclusion of extreme observations or to trimming observations at other percentiles (i.e., 1 percent, 99 percent). I also investigate the sensitivity of my results to the length of the prior period intangible return period by shortening the prior period from 48 months to 24 months. Results with a shorter prior period are consistent with the primary results. Results are slightly sensitive to the inclusion of stocks trading under $5, although *PPIR* still has a material effect on the economic and statistical significance of accruals.[26] Given that returns of low-priced stocks are affected by bid-ask bounce (which directly affects calculation of intangible returns), and low-priced stocks have characteristics distinct from stocks trading at more than $5/share (e.g., institutional attention, liquidity, short-sale constraints), it is unclear whether excluding low-priced stocks from the primary sample materially affects the robustness of the main results.

I also examine whether intangible returns proxy for investments in R&D and advertising. Prior studies suggest that investors are misled by the conservative manner in which GAAP expenses investments in R&D (Lev and Sougiannis 1996; Eberhart et al. 2004; Penman and Zhang 2002). These studies suggest that investors underestimate the implications to future earnings of current investments in R&D, leading to positive future returns in subsequent periods. In unreported analysis, I find that firms that invest heavily in R&D (high R&D intensity) and firms that are growing their R&D investment tend to have larger, and more positive, intangible returns. However, if intangible returns proxy for unrecognized investments in such things as R&D, then high R&D firms should have *lower* future returns. However, prior R&D studies show a *positive* relation between future returns and R&D activity, suggesting that intangible returns are not simply capturing a dynamic directly attributable to accounting conservatism—see Donelson and Resutek (2010) for detailed analysis of R&D investment and intangible returns.

VI. CONCLUDING REMARKS AND IMPLICATIONS FOR FUTURE RESEARCH

This study re-examines the evidence that accruals are negatively associated with future returns. Using a three-period empirical model, I examine the relation between future returns in period three, current period accruals in period two, and prior period firm growth and contraction in period one. I operationalize the empirical tests through a log-linear return decomposition of the book-to-market ratio. Using this empirical design, I decompose the information investors use to price equities into two components: tangible and intangible information. Tangible information refers to the value-relevant information measured by summary accounting measures. Intangible information refers to all other value relevant information not captured by summary accounting measures.

The empirical evidence suggests that current period accruals are not associated with future returns. This finding holds for multiple definitions and decompositions of accruals. Rather, this study suggests that the negative relation between future returns and current period accruals documented in prior literature is due to the relation between prior period intangible returns and accruals. Hence, once prior period intangible returns are controlled, the negative association between future returns and current period accruals disappears. The results are also consistent with the premise that the accrual anomaly is a derivative of the value/growth anomaly. This finding builds on the work of Desai et al. (2004) by showing that a common economic mechanism drives the negative future return pattern associated with the accrual anomaly and the value/growth anomaly.

[26] In unreported tests, I expand my sample to include stocks trading between $1/share and $5/share as of June 30, t+1. Inferences from the primary tables are qualitatively identical to those in the paper, although a marginal negative association exists between future returns and total accruals in some of the robustness tests specifications.

Finally, this study provides evidence on the relations among future returns, external financing activities, and accruals. Consistent with Daniel and Titman (2006), I show tangible log returns that represent accounting based performance are not associated with future returns, whereas transactions with outside equity holders and intangible log returns are negatively associated with future returns. This finding builds on the work of Dechow et al. (2008) and Cohen and Lys (2006) and related work examining the relations between current period accruals and external financing activities with future returns. These studies note that accruals are associated with changes in invested capital and these changes are associated with diminishing returns on investment. This study confirms this relation, but shows that a material amount of the information responsible for the subsequent return reversal is due to pricing dynamics occurring in the periods prior to accrual recognition and orthogonal to the information captured by summary accounting measures.

The results from this study provide a new perspective on the relation between accruals and future returns. From this perspective come many new questions, particularly relating to the composition of intangible returns and the interpretation of how accounting properties and principles affect intangible returns. These are not simple questions to answer empirically as any definitively defined empirical construct for intangible information will have to be uncorrelated with accounting information and highly correlated with contemporaneous returns. Complicating this exercise more is the fact that future research needs to consider that intangible returns may not (only) represent the mispricing of some type of information. As Daniel and Titman (2006) note, the negative relation between future returns and intangible returns is consistent with certain rational pricing models. Thus, future research must be careful to not overlook rational pricing dynamics as an explanation for the future return patterns associated with intangible returns.

Despite the nebulous composition and multiple interpretations of intangible returns, the primary finding of this study is unambiguous. Consistent with the tenets of rational markets, the empirical evidence in this study shows that the publicly available information contained in accruals is not mispriced by investors. This finding does not diminish the importance of accruals or that of other summary accounting metrics. The significant abnormal returns and share volume occurring in the short window around earnings announcements emphasizes the importance of accounting information to investors. However, this study suggests that researchers and investors need to consider non-accounting (intangible) information from prior periods when drawing inferences of how accounting information is incorporated into stock prices.

APPENDIX

All data are from the annual CRSP/Compustat Merged Database.

Book Value of Equity (B_t)

To obtain shareholder's equity I use, when not missing, stockholder's equity (data216). If data216 is missing, I use total common equity (data60) plus preferred stock − par value (data130), if both items are not missing. Otherwise, I use total assets (data6) minus total liabilities (data181), if both items are not missing. If none of these produce a value total shareholder's equity value, then it is treated as missing for that year.

To obtain book equity, I subtract from shareholder's equity (as computed above) the preferred stock value. Specifically, I subtract from shareholder's equity preferred stock − redemption value (data56). If data56 is missing, I subtract from shareholder's equity preferred stock − liquidating value (data10). If both data56 and data10 are missing, I subtract from shareholder's equity preferred stock − par value (data130). If data56, data10, and data130 are all missing, I set book value of equity as missing for this year. In addition, when not missing, I add to book value of equity deferred taxes − balance sheet (data35) and subtract the FASB No. 106 adjustment (data330).

Net Operating Assets (NOA_t)

To obtain operating assets (OA), I subtract from total assets (data6) financial assets. Financial assets (FA) are defined as the sum of Investments and Advances − Other (data32) plus short-term investments (data193) plus cash and short-term investments (data1).

To obtain operating liabilities (OL), I subtract financial liabilities from total liabilities (data181). Financial liabilities (FL) equal the sum of long-term debt (data9) plus debt in current liabilities (data34) plus notes payable (data206) plus the amount of preferred stock in shareholder's equity (as computed above). Accordingly, Net Operating Assets (NOA) is computed as $OA - OL$. Total accruals ($TACC_t$) as $\log(NOA)_t - \log(NOA)_{t-1}$. Data items 32, 34, and 193 are set to 0 when missing.

REFERENCES

Barberis, N., A. Schleifer, and R. Vishny. 1998. A model of investor sentiment. *The Journal of Finance* 49 (3): 307–343.

Beaver, W. H. 2002. Perspectives on recent capital market research. *The Accounting Review* 77 (2): 453–474.

———, M. McNichols, and R. Price. 2007. Delisting returns and their effect on accounting based market anomalies. *Journal of Accounting and Economics* 43 (2–3): 341–368.

Cohen, D. A., and T. Lys. 2006. Weighing the evidence on the relation between external corporate financing activities, accruals and stock returns. *Journal of Accounting and Economics* 42 (1–2): 87–105.

Daniel, K., and S. Titman. 2006. Market reactions to tangible and intangible information. *The Journal of Finance* 61 (4): 1605–1643.

Dechow, P. M., S. Richardson, and R. Sloan. 2008. The persistence and pricing of the cash component of earnings. *Journal of Accounting Research* 46 (3): 537–566.

Desai, H., S. Rajgopal, and M. Venkatachalam. 2004. Value-glamour and accruals mispricing: One anomaly or two? *The Accounting Review* 79 (2): 355–385.

Donelson, D., and R. Resutek. 2010. The effect of R&D on future returns and earnings forecasts. Working paper, Tuck School of Business at Dartmouth.

Eberhart, A., W. Maxwell, and A. Siddique. 2004. An examination of long-term abnormal stock returns and operating performance following R&D increases. *The Journal of Finance* 59 (2): 623–650.

Fairfield, P. M., J. Whisenant, and T. Yohn. 2003. Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review* 78 (1): 353–371.

Fama, E. F., and J. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy* 81 (3): 607.

———, and K. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47 (2): 427–465.

———, and ———. 1995. Size and book-to-market factors in earnings and returns. *The Journal of Finance* 50 (1): 131–155.

———, and ———. 2008. Average returns, B/M, and share issues. *The Journal of Finance* 63 (6): 2971–2995.

Feltham, G. A., and J. Ohlson. 1995. Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research* 11 (2): 689–731.

Hovakimian, A., T. Opler, and S. Titman. 2001. The debt-equity choice. *Journal of Financial and Quantitative Analysis* 36 (1): 1–24.

Ikenberry, D., J. Lakonishok, and T. Vermaelen. 1995. Market underreaction to open market share repurchases. *Journal of Financial Economics* 39 (2–3): 181–208.

Khan, M. 2008. Are accruals mispriced? Evidence from tests of an intertemporal capital asset pricing model. *Journal of Accounting and Economics* 45 (1): 55–77.

Knez, P., and M. Ready. 1997. On the robustness of size and book-to-market in cross-sectional regressions. *The Journal of Finance* 52 (4): 1355–1382.

Kothari, S. P., E. Loutskina, and V. Nikolaev. 2006. Agency theory of overvalued equity as an explanation for the accrual anomaly. Working paper, Massachusetts Institute of Technology.

Kraft, A., A. Leone, and C. Wasley. 2006. An analysis of the theories and explanations offered for the mispricing of accruals and accrual components. *Journal of Accounting Research* 44 (2): 297–339.

———, ———, and ———. 2007. Regression-based tests of the market pricing of accounting numbers: The Mishkin test and ordinary least squares. *Journal of Accounting Research* 45 (5): 1081–1114.

Lakonishok, J., A. Shleifer, and R. Vishny. 1994. Contrarian investment, extrapolation, and risk. *The Journal of Finance* 49 (5): 1541–1578.

Lev, B., and T. Sougiannis. 1996. The capitalization, amortization, and value-relevance of R&D. *Journal of Accounting and Economics* 21 (1): 107–138.

Loughran, T., and J. Ritter. 1995. The new issues puzzle. *The Journal of Finance* 50 (1): 23–51.

Mashruwala, C., S. Rajgopal, and T. Shevlin. 2006. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics* 42 (1–2): 3–33.

Mishkin, F. 1983. A rational expectations approach to macroeconomics: Testing policy effectiveness and efficient markets models. Chicago, IL: University of Chicago Press.

Penman, S. H., and X. Zhang. 2002. Accounting conservatism, the quality of earnings, and stock returns. *The Accounting Review* 77 (2): 237–264.

Richardson, S. A., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437–485.

———, ———, ———, and ———. 2006. The implications of accounting distortions and growth for accruals and profitability. *The Accounting Review* 81 (3): 713–743.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Zhang, X. F. 2007. Accruals, investment, and the accrual anomaly. *The Accounting Review* 72 (5): 1333–1363.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1375–1410

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1375

Value Relevance of FAS No. 157 Fair Value Hierarchy Information and the Impact of Corporate Governance Mechanisms

Chang Joon Song
Sungkyunkwan University

Wayne B. Thomas
Han Yi
The University of Oklahoma

ABSTRACT: Statement of Financial Accounting Standards No. 157 (FAS No. 157), *Fair Value Measurements*, prioritizes the source of information used in fair value measurements into three levels: (1) Level 1 (observable inputs from quoted prices in active markets), (2) Level 2 (indirectly observable inputs from quoted prices of comparable items in active markets, identical items in inactive markets, or other market-related information), and (3) Level 3 (unobservable, firm-generated inputs). Using quarterly reports of banking firms in 2008, we find that the value relevance of Level 1 and Level 2 fair values is greater than the value relevance of Level 3 fair values. In addition, we find evidence that the value relevance of fair values (especially Level 3 fair values) is greater for firms with strong corporate governance. Overall, our results support the relevance of fair value measurements under FAS No. 157, but weaker corporate governance mechanisms may reduce the relevance of these measures.

Keywords: *FAS No. 157; fair value; fair value hierarchy; value relevance; corporate governance.*

Data Availability: *Data are publicly available from the sources identified in the study.*

We are grateful for helpful comments received from Matt Crow, Vicki Dickinson, Karen Hennes, Leslie Hodder, Ole-Kristian Hope, William Kinney, Mark Kohlbeck, Christian Leuz, Tom Linsmeier, Bharat Sarath, participants of the 2008 Conference on Financial Economics and Accounting, the 2009 International Conference on Assurance and Governance (University of Florida), the 2009 Spring Research Conference on Corporate Governance (The University of Texas at Austin), the 2009 AAA Annual Meeting, the FASB Office Hours, and two anonymous reviewers. We also appreciate the research assistance of Ken Bills. Professor Song gratefully acknowledges support from Virginia Polytechnic Institute and State University.

Editor's note: Accepted by Mark Trombley.

Submitted: February 2009
Accepted: November 2009
Published Online: June 2010

I. INTRODUCTION

In response to users' requests for a comprehensive framework for measuring and disclosing fair value, the Financial Accounting Standards Board (FASB) issued Statement of Financial Accounting Standards No. 157, *Fair Value Measurements* (FAS No. 157; FASB 2006) in 2006, effective for fiscal years beginning after November 15, 2007. FAS No. 157 does not increase the use of fair value measurements, but instead provides a uniform definition of fair value, establishes a framework for measuring fair value, and expands disclosure about fair value measurements. More specifically and of particular interest to our study, FAS No. 157 requires fair value assets and liabilities to be disclosed by levels, where levels are based on inputs used to measure fair values: (1) Level 1 (observable inputs from quoted prices in active markets), (2) Level 2 (indirectly observable inputs from quoted prices of comparable items in active markets, identical items in inactive markets, or other market-related information), and (3) Level 3 (unobservable, firm-generated inputs).[1]

First, we test the value relevance of fair value measures for each of the three disclosure levels. A long-standing debate of fair value accounting has been centered on the trade-off between relevance and reliability. Proponents of fair value accounting argue that fair value information has greater relevance, more accurately reflects real volatility, and simplifies financial reporting (e.g., hedge accounting). In contrast, opponents of fair value accounting argue that fair value measurements are less verifiable by investors, subject to greater estimation error by management, and prone to greater managerial manipulation. These shortcomings create information asymmetry between investors and managers that can be a serious threat to the reliability of fair values (Landsman 2007; Penman 2007). Because these problems are expected to become more severe as fair value inputs become less observable by investors, we are particularly interested in testing the value relevance of Level 1 fair values versus Level 3 fair values. The reliability and therefore potential information asymmetry problems associated with Level 2 fair values potentially fall between those of Level 1 and Level 3 fair values.

As our second test, we examine whether the value relevance of fair values depends on firms' corporate governance mechanisms. The motivation for this test arises from the greater subjectivity on the part of management in measuring and reporting fair values. Although in some instances managers may use their private information to credibly report fair values (e.g., Barth et al. 1998), prior studies also provide evidence that managers may manipulate inputs for fair values for their own interests (e.g., Aboody et al. 2006; Bartov et al. 2007). For firms with weaker corporate governance mechanisms, information asymmetry problems associated with fair values may be greater, leading to more severe moral hazard problems, and therefore lower value relevance of these disclosures. On the whole, Level 1 fair value reporting is likely to suffer the least from information asymmetry, and one might expect corporate governance mechanisms to have the least impact on the valuation of these fair values. Instead, corporate governance should probably play a larger role in the value relevance of Level 3 fair values where information asymmetry is likely the highest.

Using a sample of quarterly reports by banking firms in 2008, we provide the following results. First, we find evidence that fair value measurements of Level 1, Level 2, and Level 3

[1] The FASB Staff Position (FSP) FAS No. 157-4 (FASB 2009) was issued in April 2009 to clarify the application of FAS No. 157. It provides additional guidance on estimating fair values when the volume and level of activity for asset or liability have significantly decreased. Specifically, FSP FAS No. 157-4 provides a list of factors that a company should evaluate in order to determine whether a significant decrease has occurred in the volume and level of activity for the asset or liability in relation to normal market activity for the asset or liability. When estimating fair value, the company should place (1) more weight on transactions that the company concludes are orderly; (2) less weight on transactions for which the company does not have sufficient information to conclude whether the transaction is orderly; and (3) little, if any, weight on transactions that the company concludes are not orderly.

inputs are value-relevant. We also find that the valuation coefficient of Level 3 assets is significantly less positive than those of Level 1 and Level 2 assets. The lower valuation of Level 3 assets is consistent with investors decreasing the weight they place on less reliable fair value measurements (Maines and Wahlen 2006). The valuation coefficient of Level 3 liabilities is significantly more negative than those of Level 1 and Level 2 liabilities, consistent with less reliable amounts for Level 3 liabilities. We additionally test whether the coefficient on each Level is significantly different from its theoretically predicted value of 1 for assets and -1 for liabilities.[2] In particular, if investors perceive reliability concerns for Level 3 assets (liabilities), the coefficient on these fair values could be less than 1 (-1). Our results are consistent with this prediction. In order to mitigate concerns that firm size may be correlated with the types of assets held by banks and that the strength of a bank's capital ratio could be correlated with managers' choice of valuation levels, we examine whether our results hold in subsamples partitioned by firm size and Tier 1 capital ratio. We find that our results continue to hold in all subsamples.

For our second test, we examine whether the value relevance of fair values varies across six individual governance mechanisms (i.e., board independence, audit committee financial expertise, the frequency of annual audit committee meetings, the percent of shares held by institutional investors, the auditor's office size, and no material control weakness problem under Sections 302 and 404 of the Sarbanes-Oxley Act), as well as a factor score based on these mechanisms. We find that governance has a significant impact on Level 2 and Level 3 fair values. As the strength of corporate governance increases, investors' valuation of these fair value assets and liabilities increases toward the theoretically predicted coefficient values of 1 and -1. These results highlight the importance of corporate governance mechanisms in mitigating the information asymmetry problem associated with Level 2 and Level 3 inputs. Consistent with the notion that Level 1 fair value information suffers the least from information asymmetry, we find little impact of corporate governance on these fair values.

Finally, we provide two additional analyses. First, we find that the ability of Level information provided by FAS No. 157 interacted with existing asset and liability Type information has a greater explanatory power for firm value than does Type information only. This difference suggests that valuation of Level information is not simply a reordering of valuation by Type information that existed prior to FAS No. 157 (i.e., FAS No. 157 provides new information). Second, we consider how conclusions may have changed as the economic crisis worsened and market illiquidity became more severe in 2008. Although we find no evidence of a decline in the valuation of fair values during the first three quarters of 2008, we do find some evidence that the valuation of Level 3 assets and liabilities actually moved closer to their predicted values of 1 and -1. In addition, we find evidence that the impact of corporate governance for valuation of fair value assets remains significant over time.

Overall, we conclude that the fair value hierarchy required by FAS No. 157 provides useful information to investors and the strength of corporate governance appears to mitigate the information asymmetry problem arising from relatively less reliable fair value inputs. These results contribute to the literature on fair value accounting. We provide early evidence of the value relevance of new disclosures under FAS No. 157. Prior to FAS No. 157, direct tests of the association between the reliability of fair value information and equity prices were more difficult. Using the fair value hierarchy under FAS No. 157, we provide direct evidence of the value relevance of more reliable (Level 1) versus less reliable (Level 3) information.

---

[2] These theoretical predictions are based on implicit assumptions that (1) the valuation model is properly specified and (2) markets are efficient. Therefore, we do not claim that estimated coefficients must be 1 or -1. We merely use these theoretical values as benchmarks for our statistical testing.

Two recent working papers (Kolev 2009; Goh et al. 2009) provide tests of value relevance similar to ours. Although we find valuations of Level 1 and Level 2 assets and liabilities close to 1 and −1, these papers find valuations of Level 1 and Level 2 net assets significantly less than 1.[3] Furthermore, Goh et al. (2009) document that investors value Level 2 net assets less than Level 1 net assets but do not value Level 2 and Level 3 net assets differently. In contrast, we show that Level 1 and Level 2 assets are valued similarly, while Level 3 assets are valued the least. Goh et al. (2009) also document that the value relevance of net fair value assets decreases over the first three quarters of 2008, whereas we find that the value relevance of fair values does not decrease over this period.[4] Finding evidence that the value relevance of fair values does not decrease as markets become less liquid may be particularly important to standard-setters who are interested in the market's perception of the reliability of fair values during an economic crisis.

We also contribute to the literature by examining directly the association between the strength of corporate governance and value relevance of fair values. Standard-setters understand that information asymmetry would be higher for Level 3 assets and for this reason they required firms to provide additional disclosures for these items (e.g., the inputs used to measure fair value and the effect of the measurements on earnings or changes in net assets). Presumably these additional disclosures would reduce or even eliminate the information asymmetry problem. We find evidence consistent with the information asymmetry problem continuing to exist, but the strength of corporate governance appears to ameliorate this problem. These results highlight the importance of corporate governance for the value relevance of accounting information, especially for information that is potentially less reliable.

The next section describes FAS No. 157, reviews relevant prior research, and develops hypotheses. Section III outlines the sample selection process and provides descriptive statistics. Section IV presents empirical results. The study concludes with a discussion of the results.

II. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT FAS NO. 157

In recent years, U.S. Generally Accepted Accounting Principles (GAAP) have evolved toward greater use of fair values for reporting assets and liabilities. However, prior to FAS No. 157, neither a single coherent definition for fair value nor detailed guidance for applying the fair value measurement existed. In June 2003, the FASB added the fair value measurement project to its agenda to address these issues. In September 2006, the FASB released FAS No. 157, *Fair Value Measurements*. FAS No. 157 does not expand the use of fair value measurements, but provides a coherent framework for applying fair value measurements and enhances disclosures about the nature and source of fair value measurements to increase consistency and comparability.

The FASB has concluded that fair value information is relevant (e.g., see Paragraph 217 of FAS No. 133 or Paragraph C2 of FAS No. 157).[5] To the extent that established, active markets for

---

[3] We acknowledge that different samples, time periods, and variable definitions across these studies do not allow us to make direct comparisons.

[4] Other differences between our study and Kolev (2009) and Goh et al. (2009) include: (1) we have a larger sample size and/or longer sample period, (2) we consider the impact of a larger set of corporate governance mechanisms on the pricing of fair value assets under FAS No. 157, and (3) we consider the incremental pricing effect of expanded disclosure format under FAS No. 157. However, it should be noted these studies provide additional tests beyond ours that will contribute to the literature. For example, Kolev (2009) uses a changes specification, allowing tests of the value relevance of Level 3 realized and unrealized gains/losses versus net purchases and transfers.

[5] Claims have been made that fair value may not be the most relevant measurement attribute. For example, one claim is that fair value may misrepresent management's intent (e.g., intent to hold/owe an asset/a liability to its maturity). However, fair value measurements allow shareholders to assess whether holding an asset or owing a liability is appropriate by providing *timely* value change information. See Ryan (2007, Ch. 6) for other claims and related discussions. Also see Ramesh and Graziano (2004) for various discussions around the benefits and costs of historical cost versus fair value accounting.

assets/liabilities are in place, fair value is likely to be the most relevant measurement attribute. However, fair value measurements in the absence of observed prices might be unreliable due to intrinsic measurement error (noise) and management-induced error (bias). FAS No. 157 establishes a fair value hierarchy that prioritizes the inputs used to measure fair values into three broad levels, considering the relative reliability of the inputs to fair value measurements. Level 1 inputs are quoted prices in active markets, and a quoted price is a single primary basis for the fair value measurement. Therefore, the information asymmetry between preparers and users is very low for Level 1 inputs. Level 2 inputs include either observable prices in active market for comparable assets and liabilities or observable market prices in inactive markets for identical assets and liabilities. Level 2 inputs could also include prices corroborated by market-based measures (e.g., correlation with the yield curve). Level 3 inputs are not observable from the market and reflect management's assessment of the assumptions that market participants would use in pricing the asset or liability.

The key distinction between Level 2 inputs and Level 3 inputs is whether inputs are observable. This distinction is important for verifiability of fair values. Therefore, the Level 3 inputs are subject to the highest degree of information asymmetry between preparers and users. FAS No. 157 enhances fair value disclosure by requiring firms to provide fair value measurements by input levels in the hierarchy, enabling users to assess the relative reliability of fair value measurements. The Appendix provides an example of FAS No. 157 fair value measurements disclosures from a quarterly report filed by Wells Fargo & Company.

Value Relevance of Fair Value Hierarchy

Accounting information is considered to be *value-relevant* when it has the predicted association with market value of equity (Barth et al. 2001). If a significant association is found, then it is assumed that the accounting information of research interest is relevant to investors and reliable enough to be reflected in share prices. In this section, we discuss literature related to the value relevance of fair value measures. We also detail why reliability concerns could affect the value relevance of fair values. This discussion helps in formulating an expectation regarding the differential value relevance across Levels information under FAS No. 157.

Several studies document that fair values of investment securities of banks and property-liability insurers are value-relevant (Barth 1994; Petroni and Wahlen 1995; Barth et al. 1996; Eccher et al. 1996; Nelson 1996; Carroll et al. 2003). Petroni and Wahlen (1995) find that fair values for equities and Treasury securities are value-relevant, but fair values of municipal and corporate bonds are not, suggesting securities actively traded in the market are more reliably associated with market value of equity. In contrast, using a sample of closed-end mutual funds, Carroll et al. (2003) find that fair values of thinly traded securities are value-relevant.

When no established market is available (i.e., fair values are less reliable), the evidence of value relevance of other financial instruments is mixed. For example, Nelson (1996) finds that fair values of loans, deposits, and long-term debt are not value-relevant. In contrast, Barth et al. (1996) find that fair values of loans are value-relevant, whereas Eccher et al. (1996) find the value relevance of fair values of loans only in limited settings. Finally, Venkatachalam (1996) examines the value relevance of fair values of derivatives and finds that these fair values are positively associated with market values of equity.

Another stream of fair value accounting research examines how managerial opportunism affects the reliability of fair value measurements. Beaver and Venkatachalam (2003) partition U.S. bank loan fair values into three categories (i.e., non-discretionary, discretionary, and noisy components) and show that the pricing coefficient of the discretionary loan component is negative if

the managerial intent for the discretion is likely opportunistic.[6] Studies also examine whether managers use private information for inputs to option valuation models to help or hurt information users in the U.S. setting. For example, in the case of stock option expense reporting under FAS No. 123 (FASB 2004), Aboody et al. (2006) and Bartov et al. (2007) show that managerial incentives induce managers' discretion over model assumptions (e.g., the expected option life and stock volatility input assumptions), lowering option values.[7]

In addition to financial assets, several studies support the value relevance of fair value non-financial assets (e.g., Barth and Clinch 1998; Aboody et al. 1999; Easton et al. 1993; Dietrich et al. 2000). Some of these studies also examine whether the value relevance of fair value measurements varies with the reliability of the information. For example, Dietrich et al. (2000) find the reliability of fair value estimates is an increasing function of the presence of monitoring of external (as opposed to internal) appraisers. Similarly, Muller and Riedl (2002) provide evidence that market participants appear to find external appraisals to be more reliable. However, Barth and Clinch (1998) find no difference in value relevance between internal and external appraisals.

In summary, the literature generally suggests that fair value measurements are value-relevant, and that the value relevance of fair value measurements varies with the source of information. Although the source of fair value measurements is assumed to be associated with the reliability of information, direct archival tests of the association between the reliability of fair value information and equity prices were difficult prior to FAS No. 157. FAS No. 157 now explicitly categorizes the inputs to fair value measurements and allows us to assess whether the value relevance of less reliable fair values (i.e., Level 3) is different from that of more reliable fair values (i.e., Level 1).

Regarding investors and how they weigh *reported* fair values in their pricing of the firm, two factors are important to consider: the discount adjustment and the cash flow adjustment. Investors, to the extent that they perceive greater uncertainty of fair values, will adjust upward the discount rate applied to those reported amounts, resulting in less than a dollar-for-dollar valuation. Furthermore, investors, to the extent that they perceive reported assets (liabilities) to be biased upward (downward), either intentionally or unintentionally, will adjust downward for a cash flow effect. Again, this adjustment results in less than a dollar-for-dollar valuation of reported amounts. Below we discuss why investors' downward valuation adjustment is likely to be greater for reported Level 3 fair values than for reported Level 1 and Level 2 fair values. Given that our tests of value relevance estimate investors' adjustment to reported fair values (i.e., valuation coefficient on reported discounted cash flows), we do not specify whether the adjustment is more likely to be the consequence of a discount effect versus a cash flow effect. Rather, we contend that both are likely contributing factors.

Level 3 fair values are less observable and are naturally subject to greater information asymmetry between investors and management. In addition, the more subjective nature of Level 3 fair values makes them prone to greater estimation error by management (i.e., noise). Both information asymmetry and estimation error inherent to the production of specific accounting information increase investors' adverse selection, liquidity risk, and information-processing costs, all of which

---

[6] Nissim (2003) examines whether the U.S. bank managers strategically overstate bank loans, finding that the estimated degree of overstatement systematically varies with regulatory capital requirements and the change in the rate of credit losses. In contrast, Bernard et al. (1995) examine whether Danish bank managers manage fair value estimates to meet the mark-to-market based regulatory capital requirement in Denmark, but find no such evidence.

[7] In contrast, Hodder et al. (2006) find that a large proportion of firms use value-increasing discretion in their valuation model inputs. Note that FAS No. 157 does not apply to stock option compensation measurements in FAS No. 123(R). However, these studies are relevant, showing managerial discretion over input values in option value models.

increase a firm's cost of capital (Diamond and Verrecchia 1991; Baiman and Verrecchia 1996).[8] These effects hold even without the presence of moral hazard. As cost of capital increases, the value of a firm's assets decreases. From an investor's perspective, accounting amounts that are less reliable (e.g., Level 3 fair values) are assigned a higher cost of capital and, therefore, are valued less than are more reliable accounting amounts (e.g., Level 1 fair values). Thus, we expect investors' downward adjustment to Level 3 fair values to be greater because of a greater discount effect.

We now turn our discussion to cash flow effects. By cash flow effect, we mean that investors perceive management's estimate of future cash flows to differ systematically from realized future cash flows (i.e., bias). When certain accounting information is highly subjective in nature, and managers are allowed to exercise a high degree of discretion over it, managers may be more likely to generate intentional biases in their estimations (e.g., Aboody et al. 2006; Bartov et al. 2007).[9] To the extent that these biases are expected on average, investors likely adjust such estimates in valuing the firm. Specifically, if investors are concerned about possible overstatement of Level 3 fair value assets and understatement of Level 3 fair value liabilities, then they will adjust their valuation of management-reported Level 3 assets and liabilities to less than 1 and -1, respectively.

In addition, less reliable accounting information reduces the ability of investors to monitor managerial behavior, potentially reducing the firm's operating performance and future cash flows. Bushman and Smith (2001) discuss the important role of financial accounting information as a mechanism to discipline managerial behavior. As the quality of financial information deteriorates, investors lose their ability to link the activities of the manager to firm performance. In other words, without the disciplining mechanism afforded by reliable financial accounting information, managers are held less accountable for their actions and therefore operate the firm less efficiently or extract private benefits directly, both of which are detrimental to firm value.[10] Level 3 fair values are less observable, making it more difficult (compared to Level 1 fair values) for investors to link their performance to managerial decisions, reducing the efficiency of these activities.

Because of the issues discussed above, investors are likely to decrease the weight they place on less reliable Level 3 fair value measurements in their equity-pricing decisions (Maines and Wahlen 2006). Level 1 fair values are expected to suffer less from these issues, as these amounts can be easily verified by investors. Thus, we expect the pricing coefficient of Level 1 fair value measurements to be greater than that of Level 3 fair value measurements. Regarding Level 2 fair values, these measurements potentially represent the middle ground of reliability. Level 2 fair values are based on observable market inputs, but in some cases it may be more difficult for investors to observe directly how bank managers adapt those inputs to generate reported fair values. Because observability of market inputs for Level 2 fair values is meant to indicate increased reliability of fair value measurements, we expect the valuation of Level 2 fair values to be greater than that of Level 3 fair values, but we do not make a prediction for the difference in valuation of Level 1 versus Level 2 fair values. This leads to our first hypothesis (stated in alternative form):

---

[8] This concept relates to estimation risk (Barry and Brown 1985) and investors' assessment of the extent to which a firm's cash flows are correlated with other firms' cash flows (Lambert et al. 2007). The extent to which a firm's (assessed) cash flows are non-diversifiable positively impacts its cost of capital. These conjectures are supported empirically (e.g., Leuz and Verrecchia 2000; Francis et al. 2004; Hope et al. 2009).

[9] Another type of bias is unintentional optimism by managers (e.g., psychological upward bias) (Martin et al. 2006; Roll 1986).

[10] Several studies find evidence consistent with higher quality accounting information facilitating investors' ability to monitor managers' decisions and therefore improving firm performance (Lombardo and Pagano 2002; Bens and Monahan 2004; Kanodia et al. 2004; Biddle and Hilary 2006; Hope and Thomas 2008).

H1: The value relevance of Level 1 and Level 2 fair values is greater than the value relevance of Level 3 fair values.

Although not explicitly stated as part of our hypothesis, we also test whether valuation coefficients differ from their theoretically predicted values. That is, if reported fair values accurately represent their underlying economic value, then investors are expected to assign dollar-for-dollar value to these amounts. Easily verifiable in most instances, one would predict Level 1 and Level 2 assets (liabilities) to have valuation coefficients close to 1 (−1). Furthermore, consistent with arguments above to support H1, the more subjective, less reliable Level 3 assets (liabilities) would have a valuation coefficient less than 1 (−1).

The Effect of Corporate Governance on the Value Relevance of Fair Values

In the previous section, we discuss how intrinsic estimation error (noise) and management-induced error (bias) inherent in fair value measurements might reduce the ability of investors to rely on these amounts. In this section, we consider the extent to which stronger corporate governance mechanisms can mitigate these problems. We do so by testing whether the value relevance of fair values varies with the strength of a firm's corporate governance.

Given that measurement errors are likely to be more severe for inputs without observable prices (Level 3 and perhaps Level 2) than for inputs directly observable in active markets (Level 1), we expect corporate governance to be more effective at mitigating problems associated with Level 3 fair values.

Bhat (2009) provides evidence that over the 2003–2005 period market participants perceived fair value gains and losses of banks with strong corporate governance as more relevant and reliable. She interprets these results to mean that corporate governance aids market participants in evaluating the quality of fair value estimates. Aboody et al. (2006) find firms with stronger governance are less likely to understate option value estimates. Penman (2007) discusses the importance of both the competence and independence of monitors, as well as the effectiveness of internal control systems in minimizing biases in Level 3 fair value estimates.

To the extent that strong corporate governance mechanisms reduce reliability concerns relating to estimation error and management induced bias in Level 3 fair values, investors are more likely to view Level 3 fair values as relevant, which leads to our second hypothesis (stated in alternative form):

H2: Corporate governance has a greater impact on the value relevance of Level 3 fair values than on the value relevance of Level 1 or Level 2 fair values.

III. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

Sample Selection

To maximize the power of our value relevance test, we focus on the banking industry where firms have significant amounts of fair value assets and liabilities. Table 1 delineates the sample selection process. The sample firms were initially identified from Compustat Bank Fundamentals Quarterly Research File. To be included in the sample, firms need to provide the first quarter FAS No. 157 fair value hierarchy disclosure after November 15, 2007 (n = 522). To avoid the confounding effect of prices from different macroeconomic events, we focus on firms that end the first quarter on March 31, 2008, eliminating 10 non-March 31 quarter firms (n = 512). We further require firms to have price information in the Center for Research in Security Prices (CRSP) database (n = 452). To avoid the effect from extreme outliers, we follow Belsley et al. (1980) and Fox (1991), eliminating 21 observations that have studentized residuals greater than 2 in the estimation of Equation (1) below. This procedure yields an initial sample of 431 firms.

TABLE 1

Sample Selection

Process	# of Firms (1st quarter)	# of Firms (2nd quarter)	# of Firms (3rd quarter)	Total
Firms that report the first quarter FAS No. 157 information in Compustat Bank Fundamentals Quarterly after November 15, 2007.	522			
Less: Non-March 31 1st Quarter Ending Firms	(10)			
Less: Firms that do not have price information in the CRSP database	(60)			
Less: Outliers that have studentized residual greater than 2 in the estimation of Equation (1), following Belsley et al. (1980) and Fox (1991)	(21)			
Sample for the Test of H1 (Table 3)	431	419	410	1,260
Sample for the Additional Type Test (Table 7)	431			431
Less: Firms that do not have proxy statements or have incomplete information on governance variables	(26)			
Sample for the Test of H2 (Table 6)	405	398	392	1,195

This table delineates the sample selection process. The sample firms were initially identified from Compustat Bank Fundamentals Quarterly Research File. To be included in the sample, firms should provide the first quarter FAS No. 157 fair value hierarchy disclosure after November 15, 2007 (n = 522). To avoid the confounding effect of prices from different macroeconomic events, we focus on firms that end the first quarter on March 31, 2008, eliminating 10 non-March 31-quarter firms (n = 512). We further require firms to have price information in the Center for Research in Security Prices (CRSP) database (n = 452). To avoid the effect from extreme outliers, we follow Belsley et al. (1980) and Fox (1991), eliminating 21 observations that have studentized residuals greater than 2 in the estimation of Equation (1). This procedure yields an initial sample of 431 firms for H1. Hypothesis 2 requires the sample firms to have valid governance variables. We further eliminate 26 firms due to missing proxy statements or incomplete proxy statements to determine the value of corporate governance variables of interest in this study. An additional analysis (Table 7) is performed based on the first quarter observations only (n = 431) due to manual data collection of fair value asset/liability Type information.

We test H1 for the first three quarters of 2008. Some firms fall from the sample. Therefore the number of firms in the second and third quarters is reduced, resulting in a total of 1,260 firm-quarter observations for tests of H1.[11] As discussed in more detail below, tests of H2 require hand-collected corporate governance data. Thus, we collect corporate governance data only for the first quarter of 2008 (i.e., using the initial 431 firms) and assume that these governance variables remain constant in 2008. For tests of H2, missing or incomplete proxy statement data eliminate 65 bank-quarter observations for 26 firms from these tests (n = 1,195).

Descriptive Statistics

Panel A of Table 2 provides descriptive statistics on the relative size of fair value assets and liabilities from 1,260 firm-quarters. Compared to total assets and total liabilities, the mean total

[11] In sensitivity tests discussed below, we rerun all tests using a constant sample of firms (i.e., firms that have complete data for the first three quarters of 2008). None of our inferences are affected when using this slightly reduced sample.

TABLE 2
Descriptive Statistics

Panel A: Relative Size of Fair Value Assets and Liabilities

Variable	n (Firm-Quarters)	Mean	Std. Dev.	25th Percentile	50th Percentile	75th Percentile
FVA/Total Assets	1,260	14.97%	9.89%	7.93%	13.81%	19.96%
FVA1/Total Assets	1,260	1.15%	3.32%	0.00%	0.04%	0.57%
FVA2/Total Assets	1,260	13.31%	9.94%	5.93%	12.32%	18.23%
FVA3/Total Assets	1,260	0.51%	1.51%	0.00%	0.00%	0.33%
FVL/Total Liabilities	1,260	0.37%	2.57%	0.00%	0.00%	0.01%
FVL1/Total Liabilities	1,260	0.03%	0.24%	0.00%	0.00%	0.00%
FVL2/Total Liabilities	1,260	0.32%	2.36%	0.00%	0.00%	0.00%
FVL3/Total Liabilities	1,260	0.03%	0.20%	0.00%	0.00%	0.00%

Panel B: Per Share Value of Price, Non-Fair Value, Fair Value Assets and Liabilities, and Income

Variable	n (Firm-Quarters)	Mean	Std. Dev.	25th Percentile	50th Percentile	75th Percentile
PRICE	1,260	14.024	9.798	7.705	11.625	17.960
NFVA/Share	1,260	133.606	73.340	85.004	121.797	163.279
FVA1/Share	1,260	1.842	6.209	0.000	0.061	0.815
FVA2/Share	1,260	21.150	21.321	7.630	16.352	27.250
FVA3/Share	1,260	0.889	2.712	0.000	0.000	0.502
NFVL/Share	1,260	142.619	77.589	89.420	131.275	171.936
FVL1/Share	1,260	0.071	0.701	0.000	0.000	0.000
FVL2/Share	1,260	0.783	7.552	0.000	0.000	0.002
FVL12/Share	1,260	0.854	8.178	0.000	0.000	0.004
FVL3/Share	1,260	0.044	0.380	0.000	0.000	0.000
NI/Share	1,260	0.008	0.915	0.032	0.163	0.326

(continued on next page)

Panel C: Frequency of Non-Zero Reported Fair Value Assets/Liabilities by the Fair Value Hierarchy (n = 431 Firms from the First Quarter 2008)

Assets	*INVSEC*		*ADEV*		*LOAN*		*AOTHER*	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Level 1	260/431	60.32%	7/431	1.62%	11/431	2.55%	20/431	4.64%
Level 2	390/431	90.49%	65/431	15.08%	96/431	22.27%	35/431	8.12%
Level 3	122/431	28.31%	44/431	10.21%	24/431	5.57%	51/431	11.83%
Liability	*TRDL*		*LTDEBT*		*LDEV*		*LOTHER*	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Level 1	6/431	1.39%	5/431	1.16%	10/431	2.32%	11/431	2.55%
Level 2	7/431	1.62%	27/431	6.26%	85/431	19.72%	16/431	3.71%
Level 3	1/431	0.23%	6/431	1.39%	17/431	3.94%	8/431	1.86%

Panel D: Relative Frequency and Size of Fair Value Hierarchy Information within Each Fair Value Asset Type (n = Firms from the First Quarter 2008 that Have Non-Zero Asset Types)

Asset Types	*INVSEC*		*ADEV*		*LOAN*		*AOTHER*	
	Relative Freq.	Relative Size	Relative Freq.	Relative Size	Relative Freq.	Relative Size	Relative Freq.	Relative Size
Level 1	260/424	15.7%	7/108	1.2%	11/97	2.8%	20/84	1.2%
Level 2	390/424	82.2%	65/108	90.3%	96/97	62.0%	35/84	90.3%
Level 3	122/424	2.1%	44/108	8.5%	24/97	35.2%	51/84	8.5%

Panel A provides descriptive statistics on the relative size of fair value assets and liabilities by Level. *FVA* (*FVL*) indicates total fair value of assets (liabilities). *FVA1* (*FVL1*) indicates fair value of Level 1 assets (liabilities). Similarly, *FVA2*, *FVA3*, *FVL2*, and *FVL3* are fair values of Level 2 and Level 3 assets and liabilities.

Panel B provides descriptive statistics on share price, non-fair value assets (*NFVA*), non-fair value liabilities (*NFVL*), fair value assets and liabilities, and net income (all amounts per share). *FVA1* (*FVL1*) indicates fair value of Level 1 assets (liabilities). Similarly, *FVA2*, *FVA3*, *FVL2*, and *FVL3* are fair values of Level 2 and Level 3 assets and liabilities. *FVL12* is the sum of *FVL1* and *FVL2*. *NI* is income before extraordinary items for common shareholders.

Panel C provides the frequency (i.e., the number of non-zero observations) and the percentage (i.e., the frequency divided by 431 observations from the first quarter of 2008) of reported fair value assets and liabilities by both asset/liability types and the fair value hierarchy. We classify fair value assets and liabilities into the following categories: (1) *INVSEC*

(continued on next page)

= investment securities including available for sale securities and trading securities; (2) *ADEV* = derivative assets; (3) *LOAN* = loan; (4) *AOTHER* = all other asset items including mortgage service rights, federal funds sold, assets-backed security, and other investments; (5) *TRDL* = trading liabilities; (6) *LTDEBT* = debts; (7) *LDEV* = derivative liabilities; and (8) *LOTHER* = all other liability items.

Panel D provides the relative frequency and the relative size (i.e., dollar amount) of fair value hierarchy information for each fair value asset type based on the non-zero observations. For example, there are 424 non-zero investment security observations out of 431 firms from the first quarter of 2008. Within 424 firm observations, 260, 390, and 122 firms have Level 1, Level 2, and Level 3 investment securities, respectively, and the relative sizes (i.e., dollar amount) of these assets are 15.7%, 82.2% and 2.1%, respectively. We classify fair value assets into the following categories: (1) *INVSEC* = investment securities including available for sale securities and trading securities; (2) *ADEV* = derivative assets; (3) *LOAN* = loan; and (4) *AOTHER* = all other asset items including mortgage service rights, federal funds sold, assets-backed security, and other investments.

fair value assets and liabilities are about 15 percent and 0.4 percent, respectively.[12] The fair value amounts under Level 2 inputs account for most fair values.

Panel B of Table 2 presents descriptive statistics of variables used to test for the value relevance of Level information under FAS No. 157. All variables are per-share numbers. The mean share price (*PRC*) is 14.024. The mean of non-fair value assets (*NFVA*) is 133.606, and the means of fair value assets using Level 1 inputs (*FVA1*), Level 2 inputs (*FVA2*), and Level 3 inputs (*FVA3*) are 1.842, 21.150, and 0.889, respectively. Similarly, the mean of non-fair value liabilities (*NFVL*) is 142.619, whereas the means of fair value liabilities using Level 1 inputs (*FVL1*), Level 2 inputs (*FVL2*), and Level 3 inputs (*FVL3*) are 0.071, 0.783, and 0.044, respectively.

In Panel C of Table 2 we provide a matrix of the reporting frequency of fair value Levels by hand-collected Type information from the first quarter 10-Qs of our 431 sample firms. For Type information, we classify fair value assets and liabilities into the following categories: (1) investment securities (*INVSEC*) including trading securities and available-for-sale securities; (2) derivative assets (*ADEV*); (3) loans (*LOAN*); (4) other assets (*AOTHER*) including mortgage service rights, federal funds sold, asset-backed securities, and other investments; (5) trading liabilities (*TRDL*); (6) long-term debt (*LTDEBT*); (7) derivative liabilities (*LDEV*); and (8) other liabilities (*LOTHER*). Panel C shows, for example, that 260 out of 431 firms (60.32 percent) report Level 1 investment securities. Similarly, 390 firms (90.49 percent) and 122 firms (28.31 percent) report Level 2 and Level 3 investment securities, respectively. Across the four asset types, a reasonable number of firms disclose fair values at all three levels, with just a few exceptions. The average frequency across the 12 cells is 21.75 percent. However, for liabilities, the average frequency across the 12 cells is only 3.85 percent, and only one cell (Level 2 derivative liabilities, *LDEV*) has a reasonable number of observations. In sum, both the amount and the frequency of fair values are much greater for assets than for liabilities. Nevertheless, the fact that all asset/liability types are represented in all levels of FAS No. 157's hierarchy provides evidence of the potential usefulness of these expanded disclosures.

Whereas Panel C of Table 2 shows how many firms report nonzero values for each type and level, Panel D provides the relative dollar-value size of fair value hierarchy information within each fair value asset type. For example, out of 424 nonzero *INVSEC* observations from the first quarter of 2008, 260, 390, and 122 firms have Level 1, Level 2, and Level 3 investment securities, and the relative sizes of these assets are 15.7 percent, 82.2 percent, and 2.1 percent of total investment securities, respectively. The relative size is qualitatively similar to relative frequency reported in Panel C.

IV. RESULTS

Value Relevance of Fair Value Hierarchy (H1)

To test the value relevance of FAS No. 157's fair value hierarchy, we estimate the association between share prices and fair values of assets and liabilities per share using a modified Ohlson (1995) model, which has been extensively employed in the literature. Barth and Clinch (2009) provide evidence that share-deflated specifications (as opposed to equity book value-

[12] The relative size of fair value assets and non-fair value assets reported in this study is comparable to those reported in Bech and Rice (2009). They collect balance sheet data from the Consolidated Reports of Condition and Income for ensured domestic commercial banks and non-deposit trust companies. They report that in 2008 (our sample period) 81 percent of banks' assets consist of the following: loans and leases (57 percent), noninterest-earning assets (15 percent), federal funds transactions (6 percent), and other (3 percent). These amounts are typically carried at historical cost or carrying amount. The remaining 19 percent of assets are classified as investments. Of this amount, any securities *not* classified as held-to-maturity would be reported at fair value, requiring classification into Levels 1–3. Thus, the portion of fair value assets for our sample (15 percent) is similar to that of all banking institutions in the U.S.

deflated, lagged price-deflated, returns, or equity market value-deflated specifications) perform the best in reducing scale effects in the modified Ohlson (1995) model. We partition book value into non-fair value assets and liabilities and each of the fair value levels.[13]

$$PRC_{it} = \alpha_0 + \alpha_1 NFVA_{it} + \alpha_2 FVA1_{it} + \alpha_3 FVA2_{it} + \alpha_4 FVA3_{it} + \alpha_5 NFVL_{it} + \alpha_6 FVL12_{it} + \alpha_7 FVL3_{it} + \beta_1 NI_{it} + \varepsilon_{it} \quad (1)$$

Due to the low frequency of fair value liability reporting in our sample (see Table 2), we combine Level 1 and Level 2 liabilities (*FVL12*). The dependent variable, *PRC*, is per share price measured on the 10-Q filing month-end for firm i in quarter t. Other variables are as defined previously on a per share basis.

In estimating Equation (1), we pool observations from the first three quarters of 2008 for each firm. Residuals could be correlated across quarters or across firms, so we correct standard errors and related t-statistics based on two dimensions (i.e., firms and quarters) following Petersen (2009).[14] Petersen (2009) shows that this standard error adjustment (or clustering by two dimensions) produces less-biased standard errors.

Hypothesis 1 examines whether the value relevance of Level 1 and Level 2 fair values is greater than the value relevance of Level 3 fair values. In Table 3 we report results of Equation (1) and test whether the valuation coefficient is 0 (Column A), whether it is different from 1 for assets (Column B), and whether it is different from -1 for liabilities (Column C). Before reporting tests of H1, we first note that estimated coefficients for both non-fair value assets/liabilities and fair value assets/liabilities are different from 0, indicating their value relevance.[15] In addition, we find that the coefficients on *FVA1* and *FVA2* are not different from their theoretically predicted value of 1. In contrast, the coefficient on *FVA3* is significantly less than 1, meaning that investors place less weight on Level 3 fair value assets relative to Level 1 and Level 2. At the bottom of Table 3, we provide tests of whether the coefficients across Levels are equal. We find that the coefficients on Level 1 and Level 2 assets are not significantly different, and these coefficients are both significantly greater than the coefficient on Level 3 assets. These results support H1.

As for liabilities, the coefficient on *FVL12* is not statistically different from its predicted value of -1. For *FVL3*, we find the valuation coefficient to be significantly less than -1 (or greater than 1 in absolute terms) and less than the coefficient on *FVL12*. The coefficient's magnitude is consistent with investors' perception that Level 3 fair value liabilities are understated.

In summary, Level 1 and Level 2 fair value estimates are value-relevant and valuation coefficients are not statistically different from their theoretical values of 1 or -1. For Level 3 fair

[13] We believe that a price model rather than a returns model is appropriate for testing our hypotheses because our research question is to determine whether fair value hierarchy is reflected in firm value (i.e., value relevance), rather than testing whether fair value hierarchy is reflected in changes in firm value over a specific period of time (see Barth et al. 2001).

[14] As an example of how failure to adjust standard errors in this manner can affect inferences in accounting research, see Gow et al. (2010). We used *cluster2* command in STATA provided by Mitchell Petersen. Our results are robust to alternative panel data model specifications such as the firm fixed effect model.

[15] Based on discussion in Barth and Clinch (2009), our theoretical prediction is that the coefficient on non-fair value assets (liabilities) would be between 0 and 1 (-1). We note that our observed coefficients are similar to those reported in earlier research (e.g., Barth (1994) who reports a coefficient of 0.79 for historical cost book value in a regression similar to ours) and to those reported in more recent studies (e.g., Kolev (2009) reports coefficients on book value of equity between 0.709 and 0.748).

TABLE 3

Value Relevance of Fair Values Hierarchy of FAS No. 157

Dependent Variable = *Share Price*

	(A)				(B)		(C)	
Independent Variables	**Coeff.**	**Robust Std. Err.**	**t-stat Coeff. = 0**	**p-value**	**F-stat Coeff. = 1**	**p-value**	**F-stat Coeff. = −1**	**p-value**
Intercept	1.632	0.602	2.710	0.007***				
NFVA	0.801	0.080	9.990	0.000***	6.120	0.014**		
FVA1	0.968	0.102	9.490	0.000***	0.100	0.751		
FVA2	0.972	0.098	9.900	0.000***	0.080	0.773		
FVA3	0.683	0.112	6.110	0.000***	8.040	0.005***		
NFVL	−0.818	0.089	−9.170	0.000***			4.170	0.041**
FVL12	−1.006	0.134	−7.520	0.000***			0.000	0.966
FVL3	−2.185	0.291	−7.500	0.000***			16.540	0.000***
NI	2.488	1.018	2.440	0.015**				
n	1260							
Adj. R^2	56.53%							

Coefficient Comparisons	**F-stat**	**p-value**	**F-stat**	**p-value**
Test of *FVA1 = FVA2*	0.010	0.933		
Test of *FVA1 = FVA3*	17.250	0.000***		
Test of *FVA2 = FVA3*	18.630	0.000***		
Test of *FVL12 = FVL3*			11.160	0.001**

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels (two-tailed), respectively.

This table provides the result of OLS regression of share price on non-fair value and fair value assets and liabilities. The sample includes 1,260 firm-quarters of 431 distinct firms from the first three quarters of 2008. *NFVA* is non-fair value assets per share. *FVA1*, *FVA2*, and *FVA3* are fair value assets per share from Levels 1, 2, and 3 inputs, respectively. *NFVL* is non-fair value liabilities per share. *FVL12* and *FVL3* are fair value liabilities per share from combined Levels 1 and 2 and Level 3 inputs, respectively. *NI* is net income before extraordinary items per share. This table provides three sets of test statistics. Column (A) provides t-statistics testing whether the coefficient estimates are different from 0. Column (B) provides F-statistics testing whether the coefficient estimates of each Level of fair value assets are different from 1. Column (C) provides F-statistics testing whether the coefficient estimates of each Level of fair value liabilities are different from −1. Standard errors are adjusted for two dimensions (both firms and quarters) following Petersen (2009) and Wooldridge (2002).

values, valuation coefficients suggest that investors place less weight on these less reliable amounts.[16]

As a robustness test, we consider that results in Table 3 could be confounded by certain bank characteristics. For example, suppose larger banks tend to report Level 1 and Level 2 fair values, whereas smaller banks tend to report Level 3 fair values. The differences in value relevance reported in Table 3 could reflect differences in bank characteristics rather than differences in fair value levels. For similar reasons, we may also consider that differences in capital ratio are correlated with managers' choice of valuation levels. If this is the case, and investors value banks' assets according to bank characteristics, then the results in Table 3 could be confounded.

To examine these issues, we repeat the analyses with subsamples partitioned by these bank characteristics in Table 4 (i.e., large versus small banks in Panel A and high versus low Tier 1 capital ratio banks in Panel B). We find that results for H1 generally hold for all subsamples. Specifically, we find across subsamples that (1) all of fair value hierarchies are value-relevant, (2) Level 3 asset valuation is less than that of Level 1 or Level 2 assets except for the low capital ratio subsample, (3) Level 3 liability valuation is more negative than those of Level 1 and Level 2 liabilities, (4) Level 1 and Level 2 asset valuations are not significantly different from 1, (5) Level 3 asset valuation is less than 1, (6) Level 1 and Level 2 liability valuations are not significantly different from -1, and (7) Level 3 liability valuation is less than -1.[17] We also note that the valuation coefficients for small banks tend to be lower than those for large banks. These results are consistent with investors perceiving fair value estimates by managers at small banks as being less reliable.

The Effect of Corporate Governance on the Value Relevance of Fair Values (H2)

Results so far indicate that FAS No. 157 Level information is value-relevant. Given the high level of information asymmetry between managers and investors, especially relating to reliability of Level 3 and perhaps Level 2 inputs, H2 tests whether investors place differential weights on fair values across Levels, based on the firm's corporate governance mechanisms. Based on prior literature, we use six measures of corporate governance mechanisms germane to the reliability of Level 3 fair value estimates. The six governance measures include (1) board independence (*BDIND*)[18] measured by the number of independent board members divided by the number of total board members, (2) audit committee financial expertise (*ACFE*) [19] measured by the number of audit committee members with financial expertise divided by the number of total audit committee

[16] Recall from Table 1 that the number of firms in the sample decreases from 431 in the first quarter to 410 by the third quarter for tests of H1. We repeat tests using only the 410 firms that have data for all three quarters (i.e., 1,230 firm-quarter observations). We find results nearly identical to those reported in Table 3. Coefficients for *FVA1*, *FVA2*, and *FVL12* (0.958, 0.962, and -1.003, respectively) are not significantly different from 1 or -1, the coefficient on *FVA3* is less than 1 (0.611) and significant at the 0.01 level, and the coefficient on *FVL3* is less than -1 (-1.983) and significant at the 0.01 level.

[17] We also conduct tests by including firm size and capital ratio as separate independent variables in a pooled regression (untabulated). These variables are included in alternate forms (indicator, deciles, or continuous) and either interacted or non-interacted with main test variables. Under all specifications, we continue to reach the same conclusion; levels valuations are not driven by differences in firm size or capital ratio. We also consider a growth factor (i.e., the asset growth) and do not find any evidence that the control for the asset growth alters our conclusions.

[18] Dechow et al. (1996) and Beasley (1996) show that firms that are investigated by the SEC had boards with significantly lower percentages of outside members. Similarly, both Klein (2002) and Peasnell et al. (2006) find a negative relation between board independence and earnings management. Ajinkya et al. (2005) provide evidence that the percentage of outside directors on the board is negatively associated with optimistic biases in management forecasts.

[19] Section 407 of SOX required the SEC to issue a rule requiring public companies to disclose whether their audit committee has at least one financial expert. Additionally, if companies do not have financial experts on their audit committees, then they are required to disclose that fact. Recent studies provide evidence that audit committees with more financial expertise and activities are less likely to manage earnings and to issue restatements (Abbott and Parker 2000; Agrawal and Chadha 2005; Bédard et al. 2004).

TABLE 4

Value Relevance of Fair Values Hierarchy of FAS No. 157 for Large versus Small Banks and for Banks with High versus Low Capital Ratios

Panel A: Large versus Small Banks (Dependent Variable = *Share Price*)

	Large Banks				Small Banks			
Independent Variables	**Coeff.**	**Robust Std. Err.**	**t-stat Coeff. = 0**	**p-value**	**Coeff.**	**Robust Std. Err.**	**t-stat Coeff. = 0**	**p-value**
Intercept	2.579	0.934	2.760	0.006***	2.019	0.706	2.860	0.004***
NFVA	0.778	0.114	6.810	0.000***	0.770	0.099	7.780	0.000***
FVA1	1.007	0.141	7.150	0.000***	0.891	0.105	8.480	0.000***
FVA2	0.997	0.130	7.660	0.000***	0.877	0.094	9.340	0.000***
FVA3	0.681	0.145	4.680	0.000***	0.568	0.109	5.230	0.000***
NFVL	−0.798	0.125	−6.390	0.000***	−0.785	0.101	−7.790	0.000***
FVL12	−1.054	0.169	−6.230	0.000***	−1.029	0.093	−11.010	0.000***
FVL3	−2.507	0.622	−4.030	0.000***	−1.623	0.489	−3.320	0.001***
NI	3.016	0.828	3.640	0.000***	1.467	0.856	1.710	0.087*
n	631				629			
Adj. R^2	57.32%				52.66%			

	Large Banks		Small Banks	
Coefficient Comparisons	**F-stat**	**p-value**	**F-stat**	**p-value**
Test of *FVA1 = FVA2*	0.010	0.918	0.150	0.703
Test of *FVA2 = FVA3*	7.750	0.006***	28.240	0.000***
Test of *FVA1 = FVA3*	7.040	0.008***	36.560	0.000***
Test of *FVL12 = FVL3*	3.990	0.046**	1.460	0.227**

(continued on next page)

Panel B: High versus Low Tier 1 Capital Ratio Banks (Dependent Variable = *Share Price*)

Independent Variables	Banks with High Tier 1 Capital Ratio				Banks with Low Tier 1 Capital Ratio			
	Coeff.	Robust Std. Err.	t-stat Coeff. = 0	p-value	Coeff.	Robust Std. Err.	t-stat Coeff. = 0	p-value
Intercept	1.742	0.908	1.920	0.056*	0.966	0.898	1.080	0.282
NFVA	0.797	0.110	7.270	0.000***	0.794	0.122	6.510	0.000***
FVA1	0.851	0.118	7.180	0.000***	1.162	0.172	6.760	0.000***
FVA2	0.905	0.127	7.120	0.000***	1.031	0.133	7.740	0.000***
FVA3	0.489	0.132	3.700	0.000***	0.824	0.156	5.290	0.000***
NFVL	−0.802	0.119	−6.760	0.000***	−0.816	0.132	−6.180	0.000***
FVL12	−1.031	0.119	−8.690	0.000***	−1.108	0.173	−6.420	0.000***
FVL3	−2.164	0.446	−4.850	0.000***	−2.382	0.655	−3.630	0.000***
NI	3.781	1.670	2.260	0.024**	2.030	0.787	2.580	0.010***
n	576				572			
Adj. R^2	53.58%				62.50%			

Coefficient Comparisons	Banks with High Tier 1 Capital Ratio		Banks with Low Tier 1 Capital Ratio	
	F-stat	p-value	F-stat	p-value
Test of *FVA1 = FVA2*	3.510	0.061*	1.410	0.235
Test of *FVA2 = FVA3*	18.600	0.000***	1.480	0.225
Test of *FVA1 = FVA3*	15.930	0.000***	5.120	0.024**
Test of *FVL12 = FVL3*	6.780	0.009**	2.890	0.089*

*,**,*** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels (two-tailed), respectively.

Panel A provides the regression results for subsamples partitioned by the asset size of banks. Firms are classified as either large or small banks based on the median value of total assets in each quarter. See notes for Table 3 for variable definitions and the estimation methodology. The sum of observations (n = 631 + 629 = 1,260) is the same as the total number of observations reported in Table 3.

Panel B provides the regression results for subsamples partitioned by the Tier 1 capital ratio. Firms are classified to either high or low Tier 1 capital ratio banks based on the median value of Tier 1 capital ratio in each quarter. See notes for Table 3 for variable definitions and the estimation methodology. Due to missing values of Tier 1 capital ratio in Compustat, the sum of observations (n = 576 + 572 = 1,148) is lower than the total number of observations (n = 1,260) reported in Table 3.

members, (3) the frequency of annual audit committee meetings (*ACMEET*), (4) total percent of shares held by institutional investors (*INSTHOLDPCT*)[20] calculated from 13-F filings, (5) size of audit engagement office (*AUDITOFFICESIZE*),[21] and (6) no material control weaknesses problem (*NOMCW*)[22] under Sarbanes-Oxley Act (SOX) 302 or 404. Prior studies indicate that firms with independent boards, highly financially literate audit committees, active audit committees, the presence of institutional investors, audits by auditors from larger offices, and no material control weaknesses are less likely to engage in financial reporting biases.

Panel A of Table 5 provides the descriptive statistics for the six attributes of corporate governance. The mean value of *BDIND* is 0.78, reflecting the improved board independence after SOX. The mean of *ACFE* is 0.32, indicating that one-third of the audit committee members are classified as financial experts under SOX. Audit committees in our sample meet eight times annually. On average, 30 percent of shares of sample firms are held by institutional investors, and the average log of audit engagement office audit fee revenues is 15.04. Finally, 82 percent of our sample firms do not report any material control weakness problem (*NOMCW*) prior to 2007. Panel B of Table 5 provides the correlation coefficients among corporate governance variables.

To reduce the random measurement error of individual governance variables and to parsimoniously summarize the underlying latent construct of governance quality, we create a standardized governance score (*GOVSCORE*) based on the principal-component factor analysis of six aforementioned governance variables. Archival accounting research often uses this methodology to summarize investor characteristics (e.g., Bonner et al. 2003) or firm characteristics (e.g., Baik et al. 2009). The factor loadings for the varimax orthogonal rotation are shown in the first column of Panel C of Table 5, representing how individual governance variables are weighted for (i.e., correlated with) *GOVSCORE*. Consistent with our intuition, all governance variables are positively loaded in generating *GOVSCORE*. The factor analysis generates a factor with the eigenvalue of 1.878, which accounts for about 39 percent of the total variations in the original variables. The second column provides the Kaiser-Meyer-Olkin measure of sampling adequacy (Kaiser 1974). The mean of this statistic (0.62) is greater than 0.5, indicating that *GOVSCORE* well captures the underlying common factor of six individual variables (Stewart 1981). Panel D of Table 5 shows the distribution of *GOVSCORE*. Due to this standardization, the mean and the standard deviation of *GOVSCORE* are 0 and 1, respectively. Based on *GOVSCORE*, we create the decile rank (*GOVRANK*) from 0 to 9, and then scale by 9. To test H2, we estimate Equation (2).

$$\begin{aligned} PRC_{it} = {} & \alpha_0 + \alpha_1 NFVA_{it} + \alpha_2 FVA1_{it} + \alpha_3 FVA2_{it} + \alpha_4 FVA3_{it} + \alpha_{2a} FVA1_{it} * GOVRANK_i \\ & + \alpha_{3a} FVA2_{it} * GOVRANK_i + \alpha_{4a} FVA3_{it} * GOVRANK_i + \alpha_5 NFVL_{it} + \alpha_6 FVL12_{it} \\ & + \alpha_7 FVL3_{it} + \beta_1 NI_{it} + \beta_2 GOVRANK_i + \varepsilon_{it} \end{aligned} \quad (2)$$

Equation (2) is essentially identical to Equation (1) with the exception that all levels of fair

[20] Institutional investors appear to play a monitoring role over corporate financial reporting. Chung et al. (2002) find a negative association between earnings management and the presence of institutional investors. Bushee (1998) and Bange and DeBondt (1998) find that the presence of institutional investors with a long-term investment horizon reduces the level of earnings management through R&D activities. Ajinkya et al. (2005) provide evidence that institutional ownership has a favorable effect on the likelihood of forecast occurrence as well as the frequency of issuance.

[21] Previous studies use a dichotomous variable (e.g., Big N auditors) to measure the quality of the auditor. Recent studies (Francis and Yu 2009; Choi et al. 2010) further show that even within brand name auditors, the size of a specific engagement office matters for the audit quality and audit fees. Thus, following Francis and Yu (2009), we calculate the auditor office size of our sample firms based on their 2007 audit fee revenues to proxy for the quality of audit provided for our sample firms.

[22] Ashbaugh-Skaife et al. (2008) and Doyle et al. (2007) document a negative association between disclosed control weaknesses and earnings quality. Similarly, Feng et al. (2009) show that ineffective internal controls are likely to induce less accurate management earnings forecasts.

TABLE 5

Descriptive Statistics of Corporate Governance

Panel A: Descriptive Statistics

	Variable	n	Mean	Std. Dev.	25th Percentile	50th Percentile	75th Percentile
(1)	*BDIND*	405	0.78	0.12	0.71	0.80	0.88
(2)	*ACFE*	405	0.32	0.21	0.20	0.25	0.33
(3)	*ACMEET*	405	7.99	3.74	5.00	8.00	10.00
(4)	*INSTHOLDPCT*	405	0.30	0.24	0.09	0.24	0.43
(5)	*AUDITOFFICESIZE*	405	15.04	1.86	13.76	14.55	16.26
(6)	*NOMCW*	405	0.82	0.38	1.00	1.00	1.00

Panel B: Correlations

	BDIND	*ACFE*	*ACMEET*	*INST HOLD PCT*	*AUDITOFFICESIZE*	*NOMCW*
BDIND	1	−0.048	−0.061	0.076	0.001	−0.007
		0.337	**0.222**	**0.126**	**0.988**	**0.885**
ACFE	−0.109	1	0.094	0.328	0.265	0.015
	0.029		**0.060**	**<0.0001**	**<0.0001**	**0.756**
ACMEET	−0.017	0.099	1	0.164	0.207	−0.184
	0.739	**0.047**		**0.001**	**<0.0001**	**0.000**
INSTHOLDPCT	0.031	0.246	0.174	1	0.522	−0.052
	0.538	**<0.0001**	**0.000**		**<0.0001**	**0.294**
AUDITOFFICESIZE	−0.013	0.238	0.213	0.473	1	−0.149
	0.792	**<0.0001**	**<0.0001**	**<0.0001**		**0.003**
NOMCW	0.005	−0.006	−0.156	−0.068	−0.124	1
	0.922	**0.907**	**0.002**	**0.169**	**0.012**	

(continued on next page)

Panel C: Governance Factor Score Analysis and Sample Adequacy

Variables	Factor Loading Coefficients	Kaiser-Meyer-Olkin Measure of Sampling Adequacy
BODIND	0.030	0.346
ACFE	0.415	0.713
ACMEET	0.064	0.697
INSTHOLDPCT	0.468	0.592
AUDITOFFICESIZE	0.410	0.612
NOMCW	0.103	0.558
Variation Explained	39%	Mean KMO = 0.6181
Eigenvalue	1.878	

Panel D: Descriptive Statistics of Governance Factor Score and Ranking

Variable	n	Mean	Std. Dev.	Min.	25th Percentile	50th Percentile	75th Percentile	Max.
GOVSCORE	405	0.00	1.00	−2.16	−0.77	−0.22	0.58	3.64
GOVRANK	405	0.50	0.32	0.00	0.22	0.56	0.78	1.00

Panel A provides descriptive statistics of corporate governance variables. These variables include: (1) *BDIND* = board independence measured by the number of independent board members divided by the number of total board members; (2) *ACFE* = audit committee financial expertise measured by the number of audit committee members with financial expertise divided by the number of total audit committee members; (3) *ACMEET* = the frequency of annual audit committee meetings; (4) *INSTHOLDPCT* = percentage of shares held by institutional investors calculated from 13-F filings; (5) *AUDITOFFICESIZE* = log of the audit engagement office's prior year audit fee revenue following Francis and Yu (2009); and (6) *NOMCW* = an indicator of whether the firm has not disclosed any material control weakness under SOX 302 or SOX 404 in the past.

Panel B provides correlation coefficients among corporate governance variables. The upper diagonal shows Pearson correlation coefficients. The lower diagonal presents Spearman correlation coefficients. The numbers in bold are p-values.

Panel C provides the factor loading coefficients and Kaiser-Meyer-Olkin measure of sampling adequacy for our governance factor based on six governance variables of interest. Based on six governance variables, we run a principal component factor analysis, identifying a unique factor the eigenvalue of which is 1.878. The mean Kaiser-Myer-Olkin measure of sampling adequacy is greater than 0.5, indicating that six governance variables are appropriate for the factor analysis technique we employ (Stewart 1981).

Panel D provides the descriptive statistics on *GOVSCORE* and *GOVRANK*. *GOVSCORE* is the standardized governance score from the factor analysis from Panel C of Table 5 with the mean of 0 and the standard deviation of 1. *GOVRANK* is the decile rank of *GOVSCORE*, ranging from 0 to 1.

value assets are also interacted with *GOVRANK*. Because the frequency and amount of fair value assets greatly exceeds those of fair value liabilities, we limit our design to fair value assets. If investors consider that high-quality corporate governance can mitigate the concerns of managers' opportunistic reporting or errors inherent to the estimation, then we predict the coefficients on the interaction with *GOVRANK* to be positive.

Panel A of Table 6 provides the results of the estimation of Equation (2). The coefficient on fair value assets alone (i.e., *FVA1*, *FVA2*, and *FVA3*) can be interpreted as the valuation of fair value assets for firms in the lowest governance decile. The coefficients of the interaction terms are the *incremental* valuations when moving from the lowest decile to the highest decile. As shown in the table, the coefficients on all three non-interaction terms are positive, but only *FVA1* and *FVA2* are significantly different from zero. This difference suggests that Level 1 and Level 2 fair values of low governance firms are value-relevant, but Level 3 fair values are not. It is also noteworthy that the coefficients on Level 1 and Level 2 assets (0.81 and 0.83) are less than 1 at marginal significance levels (0.07 level, one-tailed).

The interaction terms are positive for all three levels, but only Level 2 and Level 3 are significant.[23] By adding the coefficient of the interaction term (e.g., *FVA1 * GOVRANK*) with that of the non-interaction term (e.g., *FVA1*), we can determine the valuation coefficient for the highest decile (i.e., firms with strongest corporate governance). For the highest decile, the valuations of Level 1 and Level 2 assets increase to near 1 (1.01 and 0.97). Furthermore, the level of corporate governance has the greatest impact on valuation of Level 3 fair value assets. The coefficient on the lowest decile (i.e., the *FVA3* term) is 0.06; this coefficient increases by 0.76 (i.e., the *FVA3 * GOVRANK* term) to 0.82 for the highest decile. Overall, these findings support the importance of corporate governance in valuing all levels of fair values, but especially Level 3 fair values which likely represent the values with greatest information asymmetry.[24,25]

For completeness, we also examine each of the six corporate governance measures. Given that each measure in isolation provides a limited picture of a firm's overall corporate governance, we view these results as exploratory for the interested reader. We find that the coefficient on *FVA2 * GOVRANK* is positive for five of our six corporate governance measures and significant for three (audit committee meetings, institutional holding, and no material control weaknesses problems). The coefficient on *FVA3 * GOVRANK* is also positive for five of the six governance measures and significant for three (board independence, auditor office size, and no material control weaknesses problem). The coefficient on *FVA1 * GOVRANK* is positive only for four of the governance measures and significant for two (audit committee meetings and no material control weaknesses

[23] It should be noted that while the coefficient for *FVA2 * GOVRANK* is significant and the coefficient for *FVA1 * GOVRANK* is not, the two coefficients have similar magnitudes and are not significantly different. Thus, we do not rule out the possibility that the insignificance of *FVA1 * GOVRANK* is due to lower power tests of fewer Level 1 assets being reported (see Table 2). However, we believe the results are primarily attributable to Level 1 assets being less affected by governance.

[24] When we restrict the sample to the 392 firms that have data for all three quarters, we find similar results. Specifically, for tests of H2, the coefficients on *FVA2 * GOVRANK* and *FVA3 * GOVRANK* are positive (0.140 and 0.441) and significant at the 0.05 level, while the coefficient on *FVA1 * GOVRANK* is positive (0.194) but not significant.

[25] As a sensitivity test, we examine the impact of corporate governance on the valuation of fair value liabilities. As discussed previously, results related to fair value liabilities should be considered exploratory. The low frequency of reported fair value liabilities (see Table 2) makes it difficult to draw reliable conclusions. Nevertheless, we interact *GOVRANK* with *FVL12* and *FVL3* in Equation (4) to test H2. First, we note that none of our inferences regarding the impact of corporate governance on fair value assets changes. Second, we find no impact of *GOVRANK* on *FVL12*. Third, we find that the coefficient on *FVL3 * GOVRANK* is significantly negative at the 0.10 level. The unexpected sign on *FVL3 * GOVRANK* is likely due to the high correlation between *FVA3 * GOVRANK* and *FVL3 * GOVRANK*. The large number of observations with zero *FVL3* reduces the ability to draw reliable conclusion regarding the coefficient on *FVL3 * GOVRANK*.

TABLE 6

The Impact of Corporate Governance Factors on the Value Relevance of Fair Value Hierarchy Information

Panel A: Regressions using Governance Factor Rank (Dependent Variable = *Share Price*)

Independent Variables	Coeff.	t-stat Coeff. = 0	p-value
Intercept	2.84	4.37	0.00***
NFVA	0.72	8.03	0.00***
FVA1	0.81	5.86	0.00***
FVA2	0.83	7.35	0.00***
FVA3	0.06	0.22	0.83
*FVA1 * GOVRANK*	0.20	0.75	0.45
*FVA2 * GOVRANK*	0.14	3.14	0.00***
*FVA3 * GOVRANK*	0.76	2.25	0.02**
NFVL	−0.74	−7.53	0.00***
FVL12	−1.03	−7.67	0.00***
FVL3	−2.22	−7.63	0.00***
NI	2.64	2.36	0.02**
GOVRANK	−0.43	−0.40	0.69
n = 1,195			
R^2 = 58.40%			

Panel B: Regressions using Specific Governance Measures (Dependent Variable = *Share Price*)

	Specific Governance Variable of Interest					
	Board Independence		Audit Committee Financial Expertise		Audit Committee Meeting Frequency	
Independent Variables	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Intercept	2.18	2.45**	1.92	2.43**	3.23	4.03***
NFVA	0.79	9.17***	0.80	9.66***	0.74	7.64***
FVA1	0.97	6.92***	0.88	8.47***	0.72	5.97***

(continued on next page)

Panel B: Regressions using Specific Governance Measures (Dependent Variable = *Share Price*)

	Specific Governance Variable of Interest					
	Board Independence		Audit Committee Financial Expertise		Audit Committee Meeting Frequency	
Independent Variables	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**
FVA2	0.96	8.60***	0.96	9.85***	0.88	7.10***
FVA3	0.48	3.78***	0.93	3.59***	0.29	0.69
*FVA1 * RANK*	−0.02	−0.18	0.29	0.97	0.37	3.11***
*FVA2 * RANK*	0.00	0.00	0.05	0.77	0.08	1.67*
*FVA3 * RANK*	0.65	1.71*	−0.43	−1.16	0.69	0.99
NFVL	−0.81	−8.50***	−0.82	−8.96***	−0.76	−6.99***
FVL12	−1.01	−7.67***	−1.04	−6.96***	−0.92	−5.74***
FVL3	−2.08	−6.39***	−2.18	−7.99***	−2.27	−8.15***
NI	2.55	2.31**	2.55	2.25**	2.48	2.35**
RANK	−0.90	−0.79	−0.47	−0.36	−2.48	−2.11**
n	1,195		1,195		1,195	
R^2	56.31%		56.53%		57.11%	

	Specific Governance Variable of Interest					
	% of Equity Held by Institutional Investors		Auditor Size Based on Engagement Office		Strong Internal Control (No History of Material Control Weakness)	
Independent Variables	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**
Intercept	3.94	6.04***	0.79	0.84	2.66	3.49***
NFVA	0.71	7.74***	0.76	8.53***	0.78	9.27***
FVA1	0.76	3.87***	0.94	7.84***	0.72	6.40***
FVA2	0.79	7.00***	0.94	7.51***	0.86	7.67***
FVA3	0.42	2.08**	0.15	0.85	0.46	3.47***
*FVA1 * RANK*	0.24	0.75	−0.06	−0.22	0.25	2.33**
*FVA2 * RANK*	0.23	5.08***	−0.02	−0.37	0.11	3.15***
*FVA3 * RANK*	0.29	1.19	0.84	3.03***	0.29	1.69*

(continued on next page)

Panel B: Regressions using Specific Governance Measures (Dependent Variable = *Share Price*)

	Specific Governance Variable of Interest					
	% of Equity Held by Institutional Investors		Auditor Size Based on Engagement Office		Strong Internal Control (No History of Material Control Weakness)	
Independent Variables	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**	**Coeff.**	**t-stat**
NFVL	−0.73	−7.32***	−0.77	−7.94***	−0.80	−8.58***
FVL12	−1.07	−8.13***	−0.96	−7.55***	−0.83	−5.51***
FVL3	−2.21	−6.56***	−2.25	−10.01***	−1.79	−4.38***
NI	2.54	2.33**	2.67	2.42**	2.42	2.23**
RANK	−2.18	−2.02**	2.48	1.86*	−0.98	−1.20
n	1,195		1,195		1,195	
R^2	58.81%		57.15%		57.46%	

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels (two-tailed), respectively.

Panel A provides the regression results of examining the impact of governance on the value relevance of fair value assets from 405 firms' 1,195 quarterly financial reports. The dependent variable is share price. *FVA1*, *FVA2*, and *FVA3* are fair value assets per share from Levels 1, 2, and 3 inputs, respectively. *NFVL* is non-fair value liabilities per share. *FVL12* and *FVL3* are fair value liabilities per share from combined Levels 1 and 2 and level 3 inputs, respectively. *NI* is net income before extraordinary items per share. *GOVRANK* is the decile rank ranging from 0 (weak governance) to 9 (strong governance), and then scaled by 9, based on the factor score; *GOVSCORE*, using principal-component factor analysis of six governance variables from 405 firms' 2007 proxy statements: (1) *BDIND* = board independence measured by the number of independent board members divided by the number of total board members; (2) *ACFE* = audit committee financial expertise measured by the number of audit committee members with financial expertise divided by the number of total audit committee members; (3) *ACMEET* = the frequency of annual audit committee meetings; (4) *INSTHOLDPCT* = percentage of shares held by institutional investors calculated from 13-F filings; (5) *AUDITOFFICESIZE* = log of the audit engagement office's prior year audit fee revenue following Francis and Yu (2009); and (6) *NOMCW* = an indicator of whether the firm has not disclosed any material control weakness under SOX 302 or SOX 404 in the past. Standard errors are adjusted for two dimensions (both firms and quarters) following Petersen (2009) and Wooldridge (2002).

Panel B provides the results of six regressions similar to Panel A of Table 6. Instead of using a factor score, six regressions in this panel report the impact of each governance measure on the value relevance fair value assets. Corporate governance variables include: (1) *BDIND* = board independence measured by the number of independent board members divided by the number of total board members; (2) *ACFE* = audit committee financial expertise measured by the number of audit committee members with financial expertise divided by the number of total audit committee members; (3) *ACMEET* = the frequency of annual audit committee meetings; (4) *INSTHOLDPCT* = percentage of shares held by institutional investors calculated from 13-F filings; (5) *AUDITOFFICESIZE* = log of the audit engagement office's prior year audit fee revenue following Francis and Yu (2009); and (6) *NOMCW* = an indicator of whether the firm has not disclosed any material control weakness under SOX 302 or SOX 404 in the past. All variables except *NOMCW* are scaled to deciles values, ranging from 0 (weak governance) to 9 (strong governance), and then scaled by 9. Standard errors are adjusted for two dimensions (both firms and quarters) following Petersen (2009) and Wooldridge (2002).

problem).[26] Overall, the analyses of individual measures continue to support the greater impact of governance on valuation of Level 2 and Level 3 assets.

Incremental Value Relevance of Fair Value Hierarchy

In this section, we consider whether FAS No. 157's Level information is incremental to existing Type information (defined in Panel C of Table 2). In particular, we examine whether the ability of Level information provided by FAS No. 157 interacted with existing asset and liability Type information has a greater explanatory power for firm value than does Type information alone.[27] The incremental value relevance of FAS No. 157 fair value hierarchy disclosures will depend on whether investors differentially value inputs across the three levels *and* whether firms report sufficient fair value asset and liability types (i.e., existing disclosures) where investors are not able to easily discern the level of input.

Based on manually collected fair value Type data (Panel C of Table 2) from the first quarter of 2008 (n = 431), we assess the incremental value relevance of the fair value hierarchy by comparing the explanatory powers of Equation (3) (Type information) versus Equation (4) (Type disaggregated into Level information).

$$PRC_{it} = \delta_0 + \delta_1 NFVA_{it} + \delta_2 INVSEC_{it} + \delta_3 LOAN_{it} + \delta_4 ADEV_{it} + \delta_5 AOTHER_{it} + \delta_6 NFVL_{it} + \delta_7 TRDL_{it} + \delta_8 LTDEBT_{it} + \delta_9 LDEV_{it} + \delta_{10} LOTHER_{it} + \beta_1 NI_{it} + \varepsilon_{it}. \quad (3)$$

$$PRC_{it} = \phi_0 + \phi_1 NFVA_{it} + \sum_{A=1}^{3} \eta_A INVSEC\ by\ Each\ Level_{it} + \sum_{B=1}^{3} \mu_B LOAN\ by\ Each\ Level_{it} + \sum_{C=1}^{3} \omega_C ADEV\ by\ Each\ Level_{it} + \sum_{D=1}^{3} \sigma_D AOTHER\ by\ Each\ Level_{it} + \alpha NFVL_{it} + \sum_{E=1}^{2} \kappa_E TRDL\ by\ Each\ Level_{it} + \sum_{F=1}^{2} \theta_F LTDEBT\ by\ Each\ Level_{it} + \sum_{G=1}^{2} \varsigma_G LDEV\ by\ Each\ Level_{it} + \sum_{H=1}^{2} \psi_H LOTHER\ by\ Each\ Level_{it} + \beta_1 NI_{it} + \varepsilon_{it}. \quad (4)$$

Table 7 provides results of Equations (3) and (4) in the left and the right columns, respectively. The R^2 from the estimation of Equation (3) is 67.75 percent and most coefficients have their predicted sign and are statistically significant. The R^2 from the estimation of Equation (4) is 69.6 percent. For this equation we do not interpret coefficients. Given the large number of fair value components, the relatively low number of nonzero observations in each cell, and the potential for high collinearity among variables, coefficients are likely unreliable. Note, however, that the sign and the magnitude of coefficients appear to be appropriate as long as the coefficients are statistically significant at the conventional level. Our focus is the explanatory power difference across these models. When we formally test whether these two R^2s are statistically different, the Vuong's Z-statistic (Vuong 1989) is 2.22 with a p-value less than 0.01, suggesting that when the Level

[26] Unlike Kolev (2009), we find no evidence that the value relevance of fair values varies based on audit committee financial expertise. His sample includes a small number of banks (88), financial services firms (35), and insurance companies (54). Kolev (2009) also finds that valuations of fair values are lower for firms with lower equity capital and firms that develop their mark-to-model internally.

[27] Clor-Proell and Warfield (2009) provide experimental evidence that the expanded disclosure format on fair value level information may improve non-professional investors' judgment performance.

TABLE 7

Incremental Value Relevance of Fair Value Hierarchy

Panel A: Dependent Variable = *Share Price*

Independent Variables	Coeff. Estimate	Robust Std. Err.	t-stat Coeff. = 0	p-value	Non-Zero # of obs./n
Intercept	1.417	0.686	2.070	0.039**	
NFVA	0.839	0.087	9.600	0.000***	
INVSEC	0.949	0.092	10.340	0.000***	424/431
ADEV	−0.042	0.949	−0.040	0.965	108/431
LOAN	0.828	0.124	6.690	0.000***	97/431
AOTHER	0.906	0.524	1.730	0.085*	84/431
NFVL	−0.847	0.093	−9.080	0.000***	
TRDL	0.165	0.272	0.610	0.545	7/431
LTDEBT	−1.043	0.113	−9.260	0.000***	32/431
LDEV	−0.062	0.988	−0.060	0.950	90/431
LOTHER	−0.661	0.508	−1.300	0.194	25/431
NI	5.132	0.924	5.560	0.000***	

n = 431

R^2 = 67.75%

Panel: B Dependent Variable = *Share Price*

Independent Variables	Coeff. Estimate	Robust Std. Err.	t-stat Coeff. = 0	p-value	Non-Zero # of obs./n
Intercept	1.498	0.696	2.150	0.032**	
NFVA	0.812	0.085	9.550	0.000***	
INVSEC1	0.907	0.096	9.400	0.000***	260/431
INVSEC2	0.927	0.088	10.590	0.000***	390/431
INVSEC3	0.507	0.112	4.550	0.000***	122/431
ADEV1	5.342	4.308	1.240	0.216	7/431
ADEV2	−0.429	0.913	−0.470	0.639	65/431
ADEV3	−1.220	5.603	−0.220	0.828	44/431

(continued on next page)

Panel: B Dependent Variable = *Share Price*

Independent Variables	Coeff. Estimate	Robust Std. Err.	t-stat Coeff. = 0	p-value	Non-Zero # of obs./n
LOAN1	2.863	0.320	8.940	0.000***	11/431
LOAN2	0.760	0.128	5.960	0.000***	96/431
LOAN3	0.352	0.401	0.880	0.380	24/431
AOTHER1	2.003	0.609	3.290	0.001***	20/431
AOTHER2	1.633	1.227	1.330	0.184	35/431
AOTHER3	0.514	0.759	0.680	0.498	51/431
NFVL	−0.816	0.090	−9.080	0.000***	
TRDL12	1.033	1.174	0.880	0.380	7/431
TRDL3	−342.747	368.552	−0.930	0.353	1/431
LTDEBT12	−1.016	0.099	−10.220	0.000***	29/431
LTDEBT3	−2.550	0.452	−5.640	0.000***	6/431
LDEV12	0.003	0.969	0.000	0.998	85/431
LDEV3	15.207	11.610	1.310	0.191	17/431
LOTHER12	−1.138	0.424	−2.690	0.008***	22/431
LOTHER3	16.129	75.303	0.210	0.831	8/431
NI	5.023	0.906	5.550	0.000***	

n = 431

R^2 = 69.60%

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels (two-tailed), respectively.

The dependent variable is price per share at the 10-Q filing date. The sample includes 431 banking firms that disclosed FAS No. 157 fair value hierarchy information in their first quarter 10-Q filings in 2008. Vuong's Z-statistic compares the explanatory power of two competing (non-nested) model specifications (Vuong 1989).

Panel A provides the regression of share price on Type information (i.e., asset/liability type). *NFVA* is non-fair value assets. *INVSEC* is investment securities including available for sale securities, and trading securities. *ADEV* is derivative assets. *LOAN* is loans. *AOTHER* is all other asset items including mortgage service rights, federal funds sold, assets-backed security, and other investments. *NFVL* is non-fair value liabilities. *TRDL* is trading liabilities. *LTDEBT* is debts. *LDEV* is derivative liabilities. *LOTHER* is all other liability items. Vuong's Z-statistic comparing the difference in the explanatory powers across two competing models: 2.22 (p-value = 0.01, two-sided).

Panel B provides the regression of share price on fair value estimates by both asset/liability types and the fair value hierarchy under FAS No. 157. For example, *INVSEC1*, *INVSEC2*, and *INVSEC3* are Level 1, Level 2, and Level 3 fair value estimates, respectively, for *INVSEC*. For liabilities, we combine Level 1 and Level 2 values into a single value. All variables are share-deflated. Standard errors are heteroscedasticity robust following Rogers (1993). Vuong's Z statistic comparing the difference in the explanatory powers across two competing models: 2.22 (p-value = 0.01, two-sided).

information is used together with the Type information, FAS No. 157 Level information is incrementally value-relevant. We caution that the economic significance may be small because the increase in R^2 is relatively small.[28] However, given that most assets and liabilities for our banking sample are not reported at fair values, the small increase in explanatory power should not be surprising.

Tests of H1 and H2 Over Time

Finally, we consider how conclusions may have changed as the economic crisis worsened and market illiquidity became more severe in 2008. We estimate Equation (1) for each of the first three quarters. The first (untabulated) result to note is the decrease in R^2 of the model over time. In the first quarter, the R^2 equals 67.57 percent but falls to 50.91 percent by the third quarter. This result likely reflects the greater noise in stock prices as the economic crisis deepened. Second, we find that the valuation coefficients of *FVA1* and *FVA2* are not significantly different from 1 for all three quarters. *FVL12* is approximately -1 for the first and second quarters and then becomes slightly less than -1 (i.e., absolute value greater than 1) and significant at the 0.10 level in the third quarter. Finally, the coefficient on *FVA3* moves slightly closer to 1 as the year proceeds (from 0.615 in the first quarter to 0.766 in the third quarter). Similarly, *FVL3* moves closer to -1 (-2.302 in the first quarter to -1.669 in the third quarter). Thus, if anything, valuations of Level 3 assets and liabilities moved closer to their theoretically predicted values of 1 and -1 as the crisis worsened.[29] Although these results provide some initial evidence of investors' perceptions of fair values during an economic crisis, we caution that these results should be viewed as exploratory at best because we do not have strong predictions on the outcome of this test.

For tests of H2, the second quarter seems to be the quarter that best distinguishes the impact of corporate governance on the valuation of fair value assets. Only in this quarter is the valuation of all three Levels significantly greater for strong corporate governance firms compared to low governance firms. Given that the economic crisis seemed to steadily worsen as the year progressed, it is difficult to draw definitive conclusions on the joint effect of corporate governance and market liquidity on valuations of fair value assets. Clearly, this area promises to be fruitful for future research as more data become available and tests are refined.

V. CONCLUSIONS

Under FAS No. 157, firms are required to disclose fair values of asset and liability types by levels, where levels are based on inputs used to generate fair values: (1) Level 1 (observable inputs from quoted prices in active markets), (2) Level 2 (indirectly observable inputs from quoted prices of comparable items in active markets, identical items in inactive markets, or other market-related information), and (3) Level 3 (unobservable, firm-generated inputs). Thus, fair values are disclosed from most reliable (Level 1) to least reliable (Level 3), with a classification that potentially falls somewhere in the middle (Level 2). Using banking firm data from the first three quarters of 2008, we examine two important research questions related to fair value information provided by banks under FAS No. 157. First, we compare the value relevance of Level 1 and Level 2 fair values to the value relevance of Level 3 fair values. Second, we consider whether the impact of corporate governance on the value relevance of fair values is greater for Level 3 assets compared to Level 1 and Level 2 assets.

---

[28] In comparison, the R^2 in a Levels-only model (i.e., Equation (1)) for these observations is 67.57 percent. Thus, Level and Type information are incremental to one another but the increase in explanatory value is relatively small.

[29] Goh et al. (2009) conclude the opposite. For their smaller sample of banks and with a different valuation model than ours, they find that the valuation of fair value assets declines over the first three quarters of 2008.

We find the following. First, all Level information is value-relevant. Level 1 and Level 2 assets (liabilities) have valuation coefficients close to their theoretically predicted value of 1 (−1). However, Level 3 assets are valued less than 1 and less than Level 1 and Level 2 assets. Level 3 liabilities have a valuation coefficient less than −1 (i.e., absolute value greater than 1) and less than those of Level 1 and Level 2 liabilities. These results are robust to potential confounding firm characteristics such as firm size and Tier 1 capital ratios. These results suggest that investors are likely to decrease the weight they place on less reliable Level 3 fair value measurements in their equity-pricing decisions due to the information risk, inherent estimation errors, and possible reporting bias.

Second, we find evidence that the value relevance of fair value assets varies with the strength of a firm's corporate governance. For firms with low corporate governance, Level 1 and Level 2 assets are below 1 (0.81 and 0.83). These valuations are marginally different from 1 at approximately the 0.07 level (one-tailed). We find that the valuation of Level 3 assets of low governance firms is close to 0 (0.06) and not significant, suggesting no value relevance. For high governance firms, we find that the valuations of Level 1 and Level 2 assets increase to near 1 (1.01 and 0.97), and valuation of Level 3 assets increases to 0.82. The increase in asset valuations for each of the levels is consistent with strong governance reducing information asymmetry and mitigating estimation errors or reporting biases, and this is especially apparent for Level 3 assets (i.e., unobservable, firm-generated amounts) where information asymmetry is expected to be the highest.

As an additional analysis, we find that FAS No. 157 fair value hierarchy disclosure disaggregating Type information to Level information (i.e., matrix format for reporting fair values based on asset/liability type and level of input) adds some incremental value relevance to existing Type information. This finding is consistent with the prediction that FAS No. 157 provides useful incremental information for investors' equity valuations.[30]

Although we recognize that fair values may affect financial statement users in contexts other than the pricing of equity securities (Holthausen and Watts 2001), we believe that our findings should be important to U.S. and international standard-setters for understanding not only the effects of FAS No. 157 disclosures, but also how future standards (e.g., the joint IASB/FASB project on financial statement presentation) can enhance existing fair value disclosures. For example, we find that investors discount less reliable fair values possibly due to information asymmetry and moral hazard problems. However, to the extent firms have strong governance, these problems appear less severe. The impact of corporate governance on the value relevance of fair values has received limited attention in academic research, especially in the U.S.

Our study is subject to a few caveats. First, our sample firms are limited to the banking industry and represent observations during the first three quarters of 2008. Until more data are analyzed, both in the cross-section and over time, the results may not be generalizable. Second, we note that the current economic crisis existed and was worsening over our sample period. It is not yet clear to what extent fair value measures *should* be value-relevant under the framework of FAS No. 157, which requires prices to be measured based on "orderly transactions," when market liquidity is low. Although standard-setters likely hope that fair values are even more informative during an economic crisis, the low liquidity in the market makes reported fair values less observable and more subjective to measure, especially for Level 2 and Level 3 items. In additional analysis, we find little evidence that the value relevance of fair value levels worsened as the

[30] In the past, users have not always been granted their requests to have information disclosed in a matrix-type format. For example, in deliberations of segment disclosures under FAS No. 131, users lobbied for disclosure of industry segment information within each geographic segment. In the end, the FASB did not adopt the matrix format for segment reporting but instead required separate disclosure of industry segment and geographic segment information (Herrmann and Thomas 2000).

economic crisis deepened. In fact, we find some preliminary evidence that some fair values gained in value relevance.

APPENDIX
An Example of FAS No. 157 Disclosure of Wells Fargo & Company from Q1 2008

13. FAIR VALUES OF ASSETS AND LIABILITIES

Fair Value Hierarchy
Under FAS 157, we group our assets and liabilities at fair value in three levels, based on the markets in which the assets and liabilities are traded and the reliability of the assumptions used to determine fair value. These levels are:

- Level 1—Valuation is based upon quoted prices for identical instruments traded in active markets.
- Level 2—Valuation is based upon quoted prices for similar instruments in active markets, quoted prices for identical or similar instruments in markets that are not active, and model-based valuation techniques for which all significant assumptions are observable in the market.
- Level 3—Valuation is generated from model-based techniques that use significant assumptions not observable in the market. These unobservable assumptions reflect our own estimates of assumptions that market participants would use in pricing the asset or liability. Valuation techniques include use of option pricing models, discounted cash flow models, and similar techniques.

The table below presents the balances of assets and liabilities measured at fair value on a recurring basis.

(in millions)	Total	Level 1	Level 2	Level 3
Balance at March 31, 2008				
Trading assets	$ 8,893	$ 1,124	$ 7,407	$ 362
Securities available for sale	81,787	41,912	33,191	6,684(2)
Mortgages held for sale	27,927	—	26,667	1,260
Mortgage servicing rights (residential)	14,956	—	—	14,956
Other assets (1)	3,167	2,226	893	48
Total	$136,730	$45,262	$68,158	$23,310
Other liabilities (1)	$(6,235)	$(3,597)	$(2,230)	$(408)
Balance at March 31, 2007				
Trading assets	$ 6,525	$ 1,572	$ 4,599	$ 354
Securities available for sale	45,443	32,412	10,223	2,808(2)
Mortgages held for sale	25,692	—	25,692	—
Mortgage servicing rights (residential)	17,779	—	—	17,779
Other assets	538	470	58	10
Total	$ 95,977	$34,454	$40,572	$20,951

(continued on next page)

The table below presents the balances of assets and liabilities measured at fair value on a recurring basis.

(in millions)	Total	Level 1	Level 2	Level 3
Other liabilities (1)	$ (3,056)	$ (1,285)	$(1,460)	$(311)

(1) Derivatives are included in this category.

(2) Non-rated asset-backed securities collateralized by auto leases represent substantially all of this balance.

The changes in Level 3 assets and liabilities measured at fair value on a recurring basis are summarized as follows:

(in millions)	Trading assets (excluding derivatives)	Securities available for sale	Mortgage held for sale	Mortgage servicing rights (residential)	Net derivative assets and liabilities	Other liabilities (excluding derivatives)
Quarter ended March 31, 2008						
Balance, beginning of quarter	$418	$5,381	$ 146	$16,763	$ 6	$(280)
Total net gains (losses) for the quarter included in:						
Net income	(68)	(8)	(5)	(2,564)	(179)	(66)
Other comprehensive income	—	42	—	—	—	—
Purchases, sales, issuances and settlements, net	12	1,269	27	757	142	17
Net transfer into/out of Level 3	—	—	1,092(3)	—	—	—
Balance, end of quarter	$362	$6,684	$1,260	$14,956	$ (31)	$(329)
Net unrealized losses included in net income for the quarter relating to assets and liabilities held at March 31, 2008 (1)	$40 (2)	$(4)	$(5) (4)	$(1,794) (4)(5)	$(27) (4)	$(66) (4)
Quarter ended March 31, 2007						
Balance, beginning of quarter	$360	$3,447	—	$17,591	$ (68)	$(282)
Total net gains (losses) included in:						
Net income	(41)	—	—	(799)	17	(6)
Purchase, sales, issuances and settlements, net	34	(639)	—	987	—	39
Balance, end of quarter	$353	$2,808	$ —	$17,779	$ (51)	$(249)
Net losses included in net income relating to assets held at March 31, 2007 (1)	$(25) (2)	$ —		$(10) (4)(5)	$(43) (4)	$(6) (4)

(1) Represents only net losses that are due to changes in economic conditions and management's estimates of fair value and excludes changes due to the collection/realization of cash flows over time.

(2) Included in other noninterest income in the income statement.

(continued on next page)

(3) Represents mortgages held for sale that were transferred from Level 2 to Level 3 due to reduced levels of market liquidity for certain residential mortgage loans.

(4) Included in mortgage banking in the income statement.

(5) Represents total unrealized losses of $1,798 million and $11 million, net of losses of $4 million and $1 million related to sales, for first quarter 2008 and 2007, respectively.

We may be required, from time to time, to measure certain assets at fair value on a nonrecurring basis in accordance with GAAP. These adjustments to fair value usually result from application of lower-of-cost-or-market accounting or write-downs of individual assets. For assets measured at fair value on a nonrecurring basis that were still held in the balance sheet at quarter end, the following table provides the level of valuation assumptions used to determine each adjustment and the carrying value of the related individual assets or portfolios at quarter end.

		Carrying value at quarter end			Total losses for quarter ended
(in millions)	**Total**	**Level 1**	**Level 2**	**Level 3**	
March 31, 2008					
Mortgages held for sale	$1,781	$ —	$1,678	$103	$ (78)
Loans held for sale	360	—	360	—	(11)
Loans (1)	546	—	546	6	(1,297)
Private equity investments	19	16	—	3	(14)
Foreclosed assets (2)	384	—	384	—	(104)
Operating lease assets	19	—	19	—	—
					$(1,504)
March 31, 2007					
Mortgages held for sale	$5,023	$ —	$5,023	$ —	$ (66)
Loans (1)	592	—	592	—	(575)
Private equity investments	3	—	—	3	(5)
Foreclosed assets (2)	225	—	225	—	(89)
					$ (735)

(1) Represents carrying value and related write-downs of loans for which adjustments are predominantly based on the appraised value of the collateral. The carrying value of loans fully charged-off, which includes unsecured lines and loans, is zero.

(2) Represents the fair value and related losses of foreclosed real estate and other collateral owned that were measured at fair value subsequent to their initial classification as foreclosed assets.

REFERENCES

Abbott, L. J., and S. Parker. 2000. Auditor selection and audit committee characteristics. *Auditing: A Journal of Practice & Theory* 19 (2): 47–66.

Aboody, D., M. E. Barth, and R. Kasznik. 1999. Revaluations of fixed assets and future firm performance: Evidence from UK. *Journal of Accounting and Economics* 26 (1): 149–178.

———, ———, and ———. 2006. Do firms understate stock option-based compensation expense disclosed under FAS 123? *Review of Accounting Studies* 11 (4): 429–461.

Agrawal, A., and S. Chadha. 2005. Corporate governance and accounting scandals. *The Journal of Law & Economics* 48 (2): 371–406.

Ajinkya, B., S. Bhojraj, and P. Sengupta. 2005. The association between outside directors, institutional investors and the properties of management earnings forecasts. *Journal of Accounting Research* 43 (3): 343–376.

Ashbaugh-Skaife, H., D. Collins, W. Kinney, Jr., and D. R. LaFond. 2008. The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review* 83 (1): 217–250.

Baik, B., D. B. Farber, and K. Petroni. 2009. Analysts' incentives and street earnings. *Journal of Accounting Research* 47 (1): 45–69.

Baiman, S., and R. Verrecchia. 1996. The relation among capital markets, financial disclosure, production efficiency, and insider trading. *Journal of Accounting Research* 34 (1): 1–22.

Bange, M. M., and W. F. DeBondt. 1998. R&D budgets and corporate earnings targets. *Journal of Corporate Finance* 4 (2): 153–184.

Barry, C., and S. Brown. 1985. Differential information and security market equilibrium. *Journal of Financial and Quantitative Analysis* 20 (4): 407–422.

Barth, M. E. 1994. Fair value accounting: Evidence from investment securities and the market valuation of banks. *The Accounting Review* 69 (1): 1–25.

———, W. H. Beaver, and W. R. Landsman. 1996. Value-relevance of banks' fair value disclosures under FAS No. 107. *The Accounting Review* 71 (4): 513–537.

———, ———, and ———. 2001. The relevance of the value relevance literature for accounting standard-setting: Another view. *Journal of Accounting and Economics* 31 (1–3): 77–104.

———, and G. Clinch. 1998. Revalued financial tangible, and intangible assets: Associations with share prices and non-market-based value estimates. *Journal of Accounting Research* 36 (3): 199–233.

———, and ———. 2009. Scale effects in capital markets-based accounting research. *Journal of Business Finance & Accounting* 36 (3–4): 253–288.

———, W. R. Landsman, and R. J. Rendleman, Jr., 1998. Option pricing-based bond value estimates and a fundamental components approach to account for corporate debt. *The Accounting Review* 73 (1): 73–102.

Bartov, E., P. Mohanram, and D. Nissim. 2007. Managerial discretion and the economic determinations of the disclosed volatility parameter for valuing ESOs. *Review of Accounting Studies* 12 (1): 155–179.

Beasley, M. S. 1996. An empirical analysis of the relation between the board of director composition and financial statement fraud. *The Accounting Review* 71 (4): 443–465.

Beaver, W., and M. Venkatachalam. 2003. Differential pricing of components of bank loan fair values. *Journal of Accounting, Auditing & Finance* 18 (1): 41–67.

Bech, M., and T. Rice. 2009. Profits and balance sheet developments at U.S. commercial banks in 2008. *Federal Reserve Bulletin* (June): A57–A97.

Bédard, J., S. M. Chtourou, and L. Courteau. 2004. The effect of audit committee expertise, independence, and activity on aggressive earnings management. *Auditing: A Journal of Practice & Theory* 23 (2): 13–35.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: John Wiley & Sons, Inc.

Bens, D. A., and S. J. Monahan. 2004. Disclosure quality and the excess value of diversification. *Journal of Accounting Research* 42 (4): 691–730.

Bernard, V. L., R. Merton, and K. Palepu. 1995. Mark-to-market accounting for U.S. banks and thrifts: Lessons from the Danish experience. *Journal of Accounting Research* 33 (1): 1–32.

Bhat, G. 2009. Impact of disclosure and corporate governance on the association between fair value gains and losses and stock returns in the commercial banking industry. Working paper, Washington University.

Biddle, G., and G. Hilary. 2006. Accounting quality and firm-level capital investment. *The Accounting Review* 81 (5): 963–982.

Bonner, S. E., B. R. Walther, and S. M. Young. 2003. Sophistication-related differences in investors' models of the relative accuracy of analysts' forecast revisions. *The Accounting Review* 78 (3): 679–706.

Bushee, B. 1998. The Influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* 73 (3): 305–333.

Bushman, R. M., and A. J. Smith. 2001. Financial accounting information and corporate governance. *Journal of Accounting and Economics* 32 (1–3): 237–333.

Carroll, T. J., T. J. Linsmeier, and K. R. Petroni. 2003. The reliability of fair value versus historical cost information: Evidence from closed-end mutual funds. *Journal of Accounting, Auditing & Finance* 18 (1): 1–23.

Choi, J. H. J. B. Kim, C. F. Kim, and Y. Zang. 2010. Audit office size, audit quality and audit pricing. *Auditing: A Journal of Practice & Theory* 29 (1): 73–97.

Chung, R., M. Firth, and J. B. Kim. 2002. Institutional monitoring and opportunistic earnings management. *Journal of Corporate Finance* 8 (1): 29–48.

Clor-Proell, S. M., and T. D. Warfield. 2009. Financial statement presentation and nonprofessional investors' interpretation of fair value information. Working paper, University of Wisconsin–Madison.

Dechow, P., R. Sloan, and A. Sweeney. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (1): 1–36.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *The Journal of Finance* 46 (4): 1325–1359.

Dietrich, J. R., M. S. Harris, and K. A. Muller, III., 2000. The reliability of investment property fair value estimates. *Journal of Accounting and Economics* 30 (2): 125–158.

Doyle, J. T., W. Ge, and S. McVay. 2007. Accruals quality and internal control over financial reporting. *The Accounting Review* 82 (5): 1141–1170.

Easton, P., P. Eddey, and T. Harris. 1993. An investigation of revaluations of tangible long-lived assets. *Journal of Accounting Research* 31 (Supplement): 1–38.

Eccher, E. A., K. Ramesh, and S. R. Thiagarajan. 1996. Fair value disclosures by bank holding companies. *Journal of Accounting and Economics* 22 (1–3): 79–117.

Feng, M., C. Li, and S. McVay. 2009. Internal control and management guidance. *Journal of Accounting and Economics* 48 (2–3): 190–209.

Financial Accounting Standards Board (FASB). 2004. *Share-Based Payment.* Statement of Financial Accounting Standards No. 123 (Revised 2004). Norwalk, CT: FASB.

———. 2006. *Fair Value Measurements.* Statement of Financial Accounting Standards No. 157. Norwalk, CT: FASB.

———. 2009. *Determining Fair Value When the Volume and Level of Activity for the Asset or Liability Have Significantly Decreased and Identifying Transactions That Are Not Orderly.* FASB Staff Position (FSP) No. 157-4. Norwalk, CT: FASB.

Fox, J. 1991. *Regression Diagnostics.* Newbury Park, CA: Sage.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *The Accounting Review* 79 (4): 967–1010.

———, and M. Yu. 2009. The effect of Big Four office size on audit quality. *The Accounting Review* 84 (5): 1521–1552.

Goh, B., J. Ng, and K. Yong. 2009. Market pricing of banks' fair value assets reported under SFAS 157 during the 2008 economic crisis. Working paper, Singapore Management University.

Gow, I. D., G. Ormazabal, and D. J. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.

Herrmann, D., and W. B. Thomas. 2000. An analysis of segment disclosures under SFAS No. 131 and SFAS No. 14. *Accounting Horizons* 14 (3): 287–302.

Hodder, L., W. Mayew, M. McAnally, and C. Weaver. 2006. Employee stock option fair-value estimates: Do managerial discretion and incentives explain accuracy? *Contemporary Accounting Research* 23 (4): 933–975.

Holthausen, R., and R. Watts. 2001. The relevance of the value-relevance literature for financial accounting standard-setting. *Journal of Accounting and Economics* 31 (1–3): 3–75.

Hope, O.-K., and W. B. Thomas. 2008. Managerial empire building and firm disclosure. *Journal of Accounting Research* 46 (3): 591–626.

———, T. Kang, W. B. Thomas, and F. Vasvari. 2009. The effects of SFAS 131 geographic segment disclosures by U.S. multinational companies on the valuation of foreign earnings. *Journal of International Business Studies* 40 (3): 421–433.

Kaiser, H. F. 1974. An index of factorial simplicity. *Psychometrika* 39 (1): 31–36.

Kanodia, C., H. Sapra, and R. Venugopalan. 2004. Should intangibles be measured: What are the economic trade-offs? *Journal of Accounting Research* 42 (1): 89–120.

Klein, A. 2002. Audit committee, board of director characteristics, and earnings management. *Journal of*

Accounting and Economics 33 (3): 375–400.

Kolev, K. 2009. Do Investors perceive marking-to-model as marking-to-myth? Early evidence from FAS No. 157 disclosure. Working paper, New York University.

Lambert, R., C. Leuz, and R. Verrecchia. 2007. Accounting information, disclosure and the cost of capital. *Journal of Accounting Research* 45 (Supplement): 385–420.

Landsman, W. R. 2007. Is fair value accounting information relevant and reliable? Evidence from capital market research. *Accounting and Business Research* (Special Issue: International Accounting Policy Forum): 19–30.

Leuz, C., and R. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38 (Supplement): 91–124.

Lombardo, D., and M. Pagano. 2002. Law and equity markets: A simple model. In *Corporate Governance Regimes: Convergence and Diversity*, edited by Renneboog, L., J. McCahery, P. Moerland, and T. Raaijmakers, 343–362, New York, NY: Oxford University Press.

Maines, L. A., and J. M. Wahlen. 2006. The nature of accounting information reliability: Inferences from archival and experimental research. *Accounting Horizons* 20: 399–425.

Martin, R. D., J. S. Rich, and T. J. Wilks. 2006. Auditing fair value measurements: A synthesis of relevant research. *Accounting Horizons* 20 (4): 287–303.

Muller, K. A., III, and E. J. Riedl. 2002. External monitoring of property appraisal estimates and information asymmetry. *Journal of Accounting Research* 40 (3): 865–881.

Nelson, K. 1996. Fair value accounting for commercial banks: An empirical analysis of FAS No.107. *The Accounting Review* 71 (2): 161–182.

Nissim, D. 2003. Reliability of banks' fair value disclosure for loans. *Review of Quantitative Finance and Accounting* 20 (4): 355–384.

Ohlson, J. 1995. Earnings, book value and dividends in security valuation. *Contemporary Accounting Research* 11 (2): 661–687.

Peasnell, K. V., P. F. Pope, and S. E. Young. 2006. Do outside directors limit earnings management? *Corporate Finance Review* 10 (5): 5–10.

Penman, S. H. 2007. Financial reporting quality: Is fair value a plus or a minus? *Accounting and Business Research* (Special Issue: International Accounting Policy Forum): 33–44.

Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.

Petroni, K., and J. M. Wahlen. 1995. Fair values of equity and debt securities and share prices of property-liability insurer. *The Journal of Risk and Insurance* 62 (4): 719–737.

Ramesh, K., and C. M. Graziano. 2004. *Fair Value Accounting: A status report*. Danvers, MA: Financial Executive Research Foundation.

Rogers, W. H. 1993. Regression standard errors in clustered samples. *Stata Technical Bulletin* 13: 19–23.

Roll, R. 1986. The hubris hypothesis of corporate takeovers. *The Journal of Business* 59 (2): 197–216.

Ryan, S. G. 2007. *Financial Instruments & Institutions: Accounting and Disclosure Rules*. 2nd edition. Hoboken, NJ: John Wiley & Sons, Inc.

Stewart, D. 1981. The application and misapplication of factor analysis in marketing research. *Journal of Marketing Research* 18 (1): 51–62.

Venkatachalam, M. 1996. Value-relevance of banks' derivatives disclosures. *Journal of Accounting and Economics* 38 (2): 387–418.

Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57 (2): 307–333.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1411–1440

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1411

Do MD&A Disclosures Help Users Interpret Disproportionate Inventory Increases?

Yan Sun
Saint Louis University

ABSTRACT: This study investigates whether MD&A disclosures have predictive ability for future firm performance in cases of disproportionate inventory increases. Using a sample of 568 manufacturing firms with disproportionate inventory increases, I find that the favorability of explanations for inventory changes in MD&A is positively associated with a firm's profitability and sales growth in the subsequent three years. I also find that future profitability and sales growth of firms that do not explain disproportionate inventory increases in MD&A fall between those of firms with favorable explanations and firms with unfavorable explanations. These results suggest that the existence and the favorability of MD&A inventory disclosures help users interpret disproportionate inventory increases and predict future firm performance.

Keywords: *MD&A disclosures; inventory increases; firm performance; financial statement analysis.*

Data Availability: *All data are available from public sources.*

I. INTRODUCTION

This study investigates whether disclosures in the Management Discussion and Analysis (MD&A) section of Form 10-K help users interpret disproportionate inventory increases. This study addresses two specific questions. First, do explanations for disproportionate inventory increases in MD&A have predictive ability for a firm's future performance? Second, what are the implications of a firm's decision to explain disproportionate inventory increases in MD&A for its prospects? My objective is to determine whether the existence, and the favorability, of MD&A disclosures on disproportionate inventory increases contain useful information for financial statement analysis.

Studies find that inventory is an important item in financial statement analysis. Disproportionate inventory changes are negatively associated with a firm's current and future stock returns and future earnings (Lev and Thiagarajan 1993; Abarbanell and Bushee 1997, 1998). Large inventory

I am grateful for guidance from my dissertation committee members: Richard Frankel (chair), Armando Gomes, Ronald King, and Tzachi Zach. I appreciate comments and suggestions from the editor, two anonymous reviewers, Shuping Chen, Xia Chen, Daniel Collins, Nicholas Dopuch, Thomas Fields, Christopher Jones, Raj Mashruwala, Mikhail Pevzner, Frank Selto, and Phil Shane. I benefited from the input of workshop participants at Baruch College, Boston University, Drexel University, Saint Louis University, University of Colorado at Boulder, University of Minnesota, and Washington University in St. Louis, and participants at the 2007 FARS Midyear Meeting and the 2007 AAA Annual Meeting.

Editor's note: Accepted by Mark Trombley.

Submitted: June 2008
Accepted: December 2009
Published Online: June 2010

changes are accompanied by reversals of firm performance (Thomas and Zhang 2002). Financial statement analysis textbooks suggest that a disproportionate increase in inventory is a red flag that deserves careful examination (Penman 2001; Mulford and Comiskey 1996, 2002). However, disproportionate inventory increases do not have unambiguous implications for a firm's prospects. Within the group of firms with disproportionate inventory increases, inventory changes can arise under a variety of situations. Such changes can be a result of a firm's decision to increase inventory to prepare for future sales growth, inclusion of new products in inventory, a firm's decision to keep a higher inventory level (for example, to improve customer service), a downward shift in the demand for a firm's products, or earnings management. Penman (2001) recommends examining firm disclosures to better understand changes in accounting ratios. This study investigates whether MD&A disclosures can help financial information users interpret disproportionate inventory increases.

Narrative disclosures in MD&A represent a potentially important channel for firms to communicate with investors. Prior studies (e.g., Bryan 1997; Cole and Jones 2004) find that MD&A is a source of useful information for financial statement analysis. Because any known trends or events that have affected or will affect the results of operations, liquidity, and capital resources should be identified and described in MD&A (Regulation S-K), and because disproportionate inventory increases are a result of operations and can be related to liquidity, firms are likely to discuss such inventory changes in MD&A.

I identify 568 firm-years with disproportionate inventory increases in the manufacturing industry between 1998 and 2002. I read the MD&A section of Form 10-K for each firm and code the existence, and the favorability, of explanations for inventory changes. Explanations are classified as favorable if they are associated with an expectation of higher future sales. Explanations are unfavorable if they are related to lower than expected sales in the current period. To control for the possibility that the favorability of MD&A inventory disclosures is correlated with the favorability of general forward-looking information in Form 10-K, I also identify statements on a firm's performance outlook and assess the overall favorability of these statements. A firm's forward-looking statements are favorable if they suggest that profitability will remain strong or even increase, and unfavorable if they indicate that profit will be negative or lower. Using the 568 firm-years with disproportionate inventory increases and the MD&A-inventory-disclosure variables defined for each firm, I investigate the predictive ability of explanations for inventory changes in MD&A for future financial performance and the implications of a firm's disclosure decision for its prospects.

Among the 568 manufacturing firm-years with disproportionate inventory increases during 1998–2002, 282 (50 percent) provide an explanation in MD&A. Among these 282 firm-years, 119 (42 percent) have favorable explanations, 105 (37 percent) provide neutral explanations, and 58 (21 percent) have unfavorable explanations. I find that the favorability of the provided explanations is positively associated with a firm's return on assets and sales growth in the subsequent three years, suggesting that MD&A disclosures on disproportionate inventory increases, if provided, supplement financial statement information and help users predict a firm's prospects. I also find that firms not explaining the inventory changes in MD&A (hereafter non-disclosure firms) do not have significantly different future performance from disclosure firms, indicating that the set of non-disclosure firms is a pool of firms with favorable private information and firms with negative development in business or low-quality earnings. I further find that future profitability and sales growth of non-disclosure firms are greater than those of firms with unfavorable explanations, while they are lower than those of firms with favorable explanations. Taken together, the results suggest that the existence of, and the favorability of, explanations for disproportionate inventory increases in MD&A have predictive ability for future firm performance.

This study contributes to the literature in several ways. First, this study provides evidence that MD&A disclosures on inventory are useful. The disproportionate inventory increase for a particular firm does not necessarily indicate bad news. Results suggest that MD&A inventory disclosures can help users understand the inventory change and predict future firm performance. The difference in future financial performance between firms with favorable MD&A inventory disclosures and those with unfavorable disclosures is significant. Second, this study confirms the existing result that forward-looking information in MD&A is informative (e.g., Bryan 1997), and extends the literature by documenting the magnitude of the association between the favorability of forward-looking information and future accounting outcomes. Third, this study contributes to the literature on financial statement analysis by providing evidence that context matters a lot for interpreting the meaning of the levels and changes of financial ratios, in particular those related to inventory. In this respect, it supplements the results by Bernard and Noel (1991) that, in terms of predicting future accounting outcomes, it matters whether inventory increases relate to raw material, work in process, or finished goods. Overall, this study provides practical advice to financial information users seeking to investigate inventory "red flags." It also provides insight into the relevance and reliability of MD&A disclosures.

I review prior studies and develop hypotheses in Section II, select the sample in Section III, discuss research design and results in Section IV, and conclude in Section V.

II. HYPOTHESES

This study examines MD&A disclosures on inventory increases for several reasons. First, inventory is an important financial statement line item to examine in predicting future firm performance. For example, Bernard and Noel (1991) show that unexpected changes in inventory and its components are incrementally useful in predicting future earnings and sales. Thomas and Zhang (2002) find that large inventory changes are accompanied by reversals of financial performance measures and stock returns. They further find that the negative relation between accruals and future abnormal stock returns first documented by Sloan (1996) is primarily due to inventory changes. Prior research also documents that disproportionate inventory changes, measured as the percentage change in inventory minus the percentage change in sales, have incremental value relevance over earnings (Lev and Thiagarajan 1993) and are negatively associated with one-year-ahead earnings and stock returns (Abarbanell and Bushee 1997, 1998).

Second, disproportionate inventory increases arising under different situations have different implications for a firm's prospects. Although, on average, firms with disproportionate inventory increases have poorer future performance, such an inventory change for a particular firm is not necessarily a signal of bad prospects. If a firm increases its inventory level in anticipation of higher future sales or in preparation for new product introduction, then the inventory change is likely a positive indicator of future performance. However, if an inventory increase is associated with fundamental problems with the firm's products or markets, delayed inventory write-downs, overproduction, or fictitious inventory (Mulford and Comiskey 1996, 2002), then it is an unfavorable signal of future financial performance. To evaluate a firm's inventory increase and to assess its prospects, investors can utilize information from the firm.[1]

Third, MD&A is a primary vehicle by which firms can explain disproportionate inventory increases. The Securities and Exchange Commission (SEC) requires that publicly traded firms include a narrative section called Management Discussion and Analysis in audited annual reports

[1] This study looks at inventory increases, instead of decreases, because studies show that managers have a propensity to manipulate earnings upward (Fields et al. 2001) and because manufacturing firms can artificially inflate earnings with delayed inventory write-downs or overproduction (Revsine et al. 2004; Roychowdhury 2006). Investors of manufacturers are particularly interested in getting information related to inventory increases.

because the SEC believes that financial statements and their footnotes alone may be insufficient for investors to judge the quality of earnings and to predict future firm performance (SEC 1987). The SEC requires registrants to identify and describe in MD&A any known trends or any known demands, commitments, events, or uncertainties that have affected or will affect the results of operations, liquidity, or capital resources in any material way (Regulation S-K; SEC 2003). Disproportionate inventory increases are directly related to operations in the current and the future periods and can affect a firm's liquidity. Firms are likely to explain such inventory increases in MD&A, even if they make similar disclosures elsewhere.[2]

Prior research finds that MD&A is a source of useful information for financial statement analysis. For example, Bryan (1997) finds that MD&A discussions of future operations and planned capital expenditures are associated with one-year-ahead financial performance and investment decisions. Rogers and Grant (1997) report that MD&A is the basis of 31 percent of the information in an average analyst report, while financial statements provide only 14 percent of the information. Barron et al. (1999) document that the MD&A quality score measured by the SEC is negatively related to analysts' forecast error and dispersion in the subsequent year. Cole and Jones (2004) find that MD&A disclosures on revenue changes and future capital expenditures help users predict future revenues and future earnings for firms in the retail industry.

The purpose of this study is to investigate whether and how specific MD&A disclosures help users interpret changes in specific accounting numbers or ratios where there is particular demand for additional information. Financial statement analysis textbooks (e.g., Penman 2001; Mulford and Comiskey 1996, 2002) regard unusual changes in accounting ratios as red flags that deserve careful examination. Penman (2001) recommends examining firm disclosures to obtain a better understanding of changes in accounting ratios. In particular, Penman (2001) calls for investigation for changes in individual asset turnover ratios to search for explanations for these changes. Inventory turnover is one such ratio. In this study, I identify firms with disproportionate increases in inventory and examine whether related disclosures in MD&A help users understand these increases and predict future performance.

My first objective is to investigate whether explanations for inventory increases in MD&A, if provided, have predictive ability for future firm performance. Essentially, I attempt to evaluate the credibility of inventory disclosures in MD&A.[3] If the disclosures are credible, then firms with more favorable disclosures are likely to have better future performance. MD&A disclosures can be credible because misleading information in Form 10-K attracts litigation. The SEC has undertaken numerous enforcement actions against firms with inadequate or misleading MD&A disclosures.[4] MD&A disclosures supplement the information contained in financial statements. If MD&A disclosures are credible, then the provided explanations for inventory increases should help financial information users interpret a firm's inventory change, analyze its current performance, and assess

---

[2] Frankel et al. (2009) find that firms just missing earnings expectations have longer earnings conference calls, suggesting that conference calls are a channel through which investors and analysts get the information that they demand. To examine whether firms explain inventory changes during conference calls or earnings announcement press releases and thus choose not to repeat the explanations in annual reports, I look at inventory disclosures in those channels for sample firms in 2002. I find that the few firms that explain inventory increases in conference calls or earnings announcements also explain the changes in MD&A, suggesting that conference calls or earnings announcement are not the places where firms choose to exclusively explain disproportionate inventory increases. It is reasonable to examine MD&A for relevant explanations for such inventory changes.

[3] Several recent studies examine credibility of narrative disclosures in other disclosure channels. For example, Davis et al. (2008) find that the optimistic and pessimistic language that managers use in earnings press releases contains credible information about expected future firm performance. Concurrent studies also investigate firms' strategic use of narrative disclosures. Li (2008) finds that annual reports of firms with lower earnings are more difficult to read.

[4] See, e.g., *In the Matter of Edison Schools, Inc.*, Release No. 34-45925 (May 14, 2002); *In the Matter of Caterpillar Inc.*, Release No. 34-30532 (March 31, 1992).

its prospects. Then, for the subsample of firms that explain inventory increases in MD&A, I expect that the favorability of the provided explanations is positively associated with a firm's future financial performance.

However, just as litigation risks do not preclude all firms from manipulating accounting numbers, they may be insufficient to force firms to disclose credible information in MD&A. Firms with inventory increases associated with negative development in business and managers of firms with inventory increases due to inappropriate decisions may not want to disclose the real reasons underlying the inventory changes. For example, the SEC filed charges against the former CEO and CFO of Kmart in 2005, accusing them of misleading investors about Kmart's financial condition in the months preceding the company's bankruptcy. Kmart had a massive inventory overbuy in the summer of 2001. However, in the MD&A section of Kmart's Form 10-Q for the third quarter and nine months ending October 31, 2001, the managers attributed the increases in inventory to seasonal inventory fluctuations and actions taken to improve the company's overall in-stock position. The SEC claims that these statements materially misrepresented the company's liquidity (SEC 2005). If firms do not provide reliable explanations for inventory increases, then the relevant MD&A disclosures are not likely to help financial information users in their financial statement analysis. I test the predictive ability of MD&A inventory disclosures by examining the relation between the favorability of inventory disclosures and a firm's future financial performance against the following null hypothesis:

H1: The favorability of MD&A disclosures on inventory increases is not associated with a firm's future financial performance.

The second objective of this study is to investigate the implications of a firm's MD&A disclosure decision for its prospects. Although firms can explain disproportionate inventory increases in MD&A, firms may choose not to do that for several reasons. First, the inventory increases may be related to earnings management for some firms. Earnings management refers to a situation in which "managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company, or to influence contractual outcomes that depend on reported accounting numbers" (Healy and Wahlen 1999, 368). Information asymmetry is a necessary condition for earnings management (Dye 1988; Trueman and Titman 1988; Richardson (2000). For earnings management to be successful, "at least some users of accounting information must be either unable or unwilling to unravel completely the effects of the earnings management" (Fields et al. 2001, 260). Studies suggest that a firm's information asymmetry is negatively associated with its disclosure quality (Diamond and Verrecchia 1991; Lang and Lundholm 1996). Better disclosures reduce a firm's flexibility in earnings management and increase the probability that earnings management will be detected.[5] To conceal earnings management, firms that inflate the inventory level are less likely to disclose the underlying reasons for the inventory changes. Because earnings management reverses over time, *ceteris paribus*, these firms are likely to have lower profitability in the future.

Second, firms with inventory increases associated with negative business development may choose not to disclose the reasons for the inventory changes. Verrecchia (1983) demonstrates analytically that, in the presence of disclosure costs, only firms with favorable news above a certain threshold disclose their information to differentiate themselves from other firms and to

[5] Lobo and Zhang (2001) find a negative relation between analysts' disclosure rating and the magnitude of earnings management. Jo and Kim (2007) document a negative relation between disclosure frequency and earnings management. Hirst and Hopkins (1998) and Hunton et al. (2006) show with experiments that financial reporting transparency reduces the prevalence of earnings management and increases the detection of earnings management.

avoid undervaluation. If the majority of firms that do not explain inventory increases in MD&A are firms with low-quality earnings (i.e., earnings manipulators) or firms with bad news, then the existence of explanations can be an indicator of superior future firm performance.

Third, even firms with favorable private information may choose not to explain the inventory changes if disclosure costs are greater than disclosure benefits. Research shows that disclosures can improve the liquidity of a firm's stock, increase analyst following, and reduce the cost of capital (Healy and Palepu 2001; Core 2001; Lang and Lundholm 1996; Botosan 1997). But disclosures are costly. Besides the costs associated with information dissemination, firms face proprietary costs (Verrecchia 1983; Dye 1985, 1986). Release of proprietary information can reduce a firm's competitive advantage. For example, in cases of disproportionate inventory increases, a firm may receive private information about increased demand for a certain product and increase its stock of that product accordingly. Such a firm is better prepared for the demand shock than its competitors and can increase its market share. This advantage will be lost if the firm discloses the reason for its inventory buildup. Verrecchia (2001) and Dye (2001) discuss the cost-benefit trade-offs associated with disclosures. If disclosure costs are too high or if disclosure benefits are too low, then firms with inventory increases arising from favorable trends will not explain the inventory changes. However, these firms are likely to perform well in the future.

Because the set of firms that do not explain disproportionate inventory increases in MD&A can be a pool of firms with poor or good prospects, and because a firm's future performance is highly dependent on future shocks, it is unclear whether one should expect a relation between a firm's MD&A disclosure decision and its future performance. I test the following null hypothesis on the relation between the existence of MD&A disclosures on inventory increases and a firm's future financial performance:

H2: The existence of MD&A disclosures on inventory increases is not associated with a firm's future financial performance.

III. SAMPLE SELECTION

To address the research objectives, I need a sample of firms for which investors are likely to demand additional information on inventory. Such firms should have a material level of inventory and a disproportionate inventory increase. I start with all U.S. manufacturing firms (SIC 2000–3999) traded on NYSE/AMEX/NASDAQ in years between 1998 and 2002. I select sample firms from the manufacturing industry for two reasons. First, inventory is an important asset for manufacturers. A manufacturing firm procures materials, makes products, and sells finished goods. Each step of its operations involves inventory management and inventory valuation. Therefore, the level, and the quality, of inventory can reflect a manufacturing firm's current operating results and predict its future performance. For example, higher than expected demand in the current period can lead to a lower ending inventory level; an ending inventory with a higher probability of becoming obsolete implies lower profit margin(s) in future period(s). Second, manufacturing firms can overproduce to artificially inflate earnings. When the inventory level increases under absorption costing, the amount of fixed production cost charged to income decreases. Firms can artificially inflate income and gross margin by producing more products. The favorable income effect of large inventory increases can even offset an unfavorable income effect of a sales decrease (Revsine et al. 2004). Roychowdhury (2006) finds that firms, especially manufacturing firms, avoid reporting losses by overproduction to reduce the reported cost of goods sold. Because of these two reasons, financial information users, such as investors, of manufacturing firms are particularly interested in getting more information to analyze disproportionate inventory increases.

I require that firms in the sample satisfy the following criteria:

(1) The firm has necessary Compustat data to compute the change in days in inventory in

year t and performance measures (e.g., return on assets, sales growth) for year t+1. It also has CRSP monthly stock returns over the 12 months ending four months after the fiscal year-end of year t so that its stock is actively traded before and around the time when its Form 10-K for year t is filed with the SEC.

(2) At the end of year t−1, the firm has a book value of inventory at least 10 percent of its book value of total assets, so that inventory is a material asset for the firm.

(3) The firm does not change its fiscal year-end or engage in mergers or acquisitions in year t. The purpose of this criterion is to eliminate explanations for inventory changes unrelated to a firm's normal operations. Following Hribar and Collins (2002), mergers and acquisitions are identified using footnote #1 of the Compustat annual file.

(4) To identify firms with disproportionate inventory increases, I require that the firm have a percentage increase in days in inventory of at least 20 percent in year t.[6] Because an increase in selling activities can necessitate a corresponding increase in inventory, I define disproportionate inventory increases as those that cannot be readily justified by a firm's change in performance. I compute each firm's days in inventory, which represents the number of days it takes to sell the inventory, as 365/(cost of goods sold/inventory). Among the 5,056 observations satisfying criteria (1)–(3), 835 (16.52 percent) have an increase in days in inventory of at least 20 percent.

(5) The firm has a positive change in the inventory level. Days in inventory can increase even if the inventory level decreases if the reduction in cost of goods sold is sufficiently large. Therefore, I delete firm-years with a decrease in the inventory level.

(6) The Form 10-K filing for the firm-year is available on the SEC's EDGAR database. In addition, inventory numbers on Compustat correspond to those in Form 10-K. This sample selection process is described in Table 1.

There are 568 firm-years from 1998 to 2002 that satisfy these sample selection criteria.[7] Table 2 reports descriptive statistics for the sample. Current and future performance measures display considerable cross-sectional variation (Panel A). For example, although sales growth rate in year t (*SALESGR*) has a mean of 2.95 percent, it has a standard deviation of 27.40 percent; the first quartile of its distribution is −11.45 percent and the third quartile is 11.79 percent. Similarly, sales growth rate in year t+1 has a mean of 9.83 percent and a standard deviation of 34.50 percent. The other performance measures—return on assets (*ROA*), change in ROA (*ROA_CH*) and stock returns (*BHR*)—also have large standard deviations and interquartile ranges. These statistics suggest that disproportionate inventory increases arise under various circumstances. As a result, investors want to utilize additional information to better understand the implications of the inventory changes.

I report the sample distribution by two-digit SIC industry in Panel B and by year in Panel C. Table 2 provides no evidence that a particular segment of the manufacturing industry is overweighed in the sample. However, the number of sample observations, by itself or as a percentage of all observations satisfying sample selection criteria (1)–(3) in a particular year, decreases in 2001 and 2002. There are two possible explanations for this decrease. First, accounting scandals in 2001 and the Sarbanes-Oxley Act passed in 2002 reduce the prevalence of earnings management via inventory. Second, the economic recession starting in 2001 reduces realized and/or anticipated sales growth at firm level. As a result, fewer firms have disproportionate inventory increases associated with current or expected future sales growth.

---

[6] The results on hypotheses are robust to alternative definition of disproportionate inventory increases, such as an increase in days in inventory of at least 30 percent or falling in the highest decile for firms with the same two-digit SIC code.

[7] These 568 observations come from 487 distinct firms. Results on hypotheses are similar if I only use the first time that each firm has a disproportionate inventory increase during the sample period.

TABLE 1

Sample Selection

	Number of Observations
U. S. manufacturing firms (SIC 2000–3999) traded on NYSE/AMEX/ NASDAQ in year *t* between 1998 and 2002	13,102
Less:	
(1) Firms with missing Compustat data to compute the change in days in inventory in year *t* or financial performance measures for year *t* or year *t*+1, or missing CRSP monthly stock returns for any of the 12 months ending four months after the fiscal year-end of year *t*	(4,016)
(2) Firms with a book value of inventory less than 10 percent of the book value of total assets at the end of year *t*−1	(2,528)
(3) Firms changing the fiscal year-end or engaging in mergers or acquisitions in year *t*	(1,502)
(4) Firms with a percentage change in days in inventory of less than 20 percent in year *t*	(4,221)
(5) Firms without a positive change in inventory level in year *t*	(165)
(6) Firms without a Form 10-K for year *t* on the SEC's EDGAR database or with different inventory numbers on Compustat and in Form 10-K	(102)
Final Sample	568

To examine how firms with disproportionate inventory increases differ from other firms, I compare selected firm characteristics and performance measures between sample observations and non-sample observations in Table 3. Non-sample observations refer to the 4,386 firm-years satisfying sample selection criteria (1)–(3), but with an increase in days in inventory of less than 20 percent or without an increase in the inventory level. By construction, sample firms have a larger increase in days in inventory in year *t* (*INVDAYS_CH%*). The average change in days in inventory is −7.53 percent for non-sample firms and 45.62 percent for sample firms. The difference between these two groups of firms is 53.15 percent. Sample firms also have a higher variation in inventory changes (*Std_INVDAYS_CH%*) in the past five years. Although sample firms have a smaller average change in days in inventory in year *t*−1 (*INVDAYS_CH%lag1*), the difference in the average change between non-sample firms and sample firms (8.00 percent) is much smaller than the difference in year *t* (53.15 percent). Therefore, the disproportionate inventory increases for sample firms in year *t* cannot be explained by their smaller inventory changes in year *t*−1.

Sample firms have a smaller market value of common equity (*MV*), a greater variation in profitability and stock returns from year *t*−4 to year *t* (*StdROA*, *StdBHR*), fewer analysts following (*NUMEST*), and a smaller institutional ownership (*PERCINT*). Investors for firms with these characteristics are likely to demand more information from the firm even without disproportionate inventory increases.

Looking at performance measures, I find that in the current year, sample firms have a bigger decrease in ROA (*ROA_CH*) and a lower sales growth rate (*SALESGR*). In the subsequent three years, on average, sample firms have a smaller ROA (*ROAlead1,2,3*), a larger decrease in ROA (*ROA_CHlead1,2,3*), and lower buy-and-hold stock returns (*BHRlead1,2,3*), consistent with the finding in prior studies that disproportionate inventory increases are negatively associated with future profitability and stock returns (Abarbanell and Bushee 1997, 1998). However, sample firms also have higher future sales growth (*SALESGRlead1,2*). Therefore, it appears that the group of

TABLE 2

Descriptive Statistics for the Sample[a]

Panel A: Descriptive Statistics

Variables	n	Mean	Standard Deviation	1st Quartile	Median	3rd Quartile
Inventory Change						
INVDAYS_CH%	568	0.4562	0.2774	0.2602	0.3487	0.5553
INV_CH%	568	0.4074	0.3182	0.1804	0.3204	0.5459
Selected Characteristics						
MV	568	550.22	2375.51	27.24	75.96	333.10
BM	568	0.8072	0.7672	0.2899	0.5816	1.0572
StdROA	568	0.0986	0.1025	0.0338	0.0659	0.1278
StdBHR	565	0.6861	0.5969	0.3325	0.5099	0.8195
NUMEST	360	4.3333	4.0326	2.0000	3.0000	6.0000
PERCINT	565	0.3246	0.2552	0.1007	0.2619	0.4972
Current Performance						
ROA	568	0.0251	0.1938	−0.0058	0.0628	0.1303
ROA_CH	568	−0.0166	0.1408	−0.0680	−0.0172	0.0310
SALESGR	568	0.0295	0.2740	−0.1145	−0.0027	0.1179
BHR	568	0.0714	0.9185	−0.4130	−0.1527	0.2562
Future Performance						
ROAlead1	568	−0.0066	0.2102	−0.0507	0.0425	0.1089
ROAlead2	521	−0.0006	0.1928	−0.0436	0.0434	0.0997
ROAlead3	463	0.0080	0.1667	−0.0420	0.0467	0.1020
ROA_CHlead1	568	−0.0321	0.1453	−0.0862	−0.0234	0.0178
ROA_CHlead2	521	−0.0154	0.0897	−0.0488	−0.0094	0.0175
ROA_CHlead3	463	−0.0112	0.0655	−0.0385	−0.0083	0.0147
SALESGRlead1	568	0.0983	0.3450	−0.0749	0.0550	0.1998
SALESGRlead2	521	0.0678	0.2311	−0.0533	0.0457	0.1825
SALESGRlead3	462	0.0553	0.2001	−0.0508	0.0367	0.1329
BHRlead1	567	0.1808	1.0279	−0.3945	−0.0301	0.4527
BHRlead2	552	0.0455	0.4750	−0.2568	0.0125	0.2705
BHRlead3	487	0.0437	0.3542	−0.1724	0.0316	0.2274

(continued on next page)

Panel B: Sample Distribution by Two-Digit SIC Industry

Two-Digit SIC	(1) Number of Sample Observations	(2) Number of All Observations[b]	(1)/(2)
20	18	250	7%
21	4	15	27%
22	5	70	7%
23	15	169	9%
24	15	57	26%
25	11	84	13%
26	8	60	13%
27	3	64	5%
28	58	519	11%
29	4	41	10%
30	9	153	6%
31	9	88	10%
32	6	51	12%
33	23	216	11%
34	16	210	8%
35	94	788	12%
36	111	898	12%
37	29	242	12%
38	108	834	13%
39	22	145	15%
Total	568	4,954	11%

Panel C: Sample Distribution by Year

Year	(1) Number of Sample Observations	(2) Number of All Observations	(1)/(2)
1998	150	1,069	14%
1999	142	1,059	13%
2000	132	976	14%
2001	84	950	9%

(continued on next page)

Panel C: Sample Distribution by Year

Year	(1) Number of Sample Observations	(2) Number of All Observations	(1)/(2)
2002	60	900	7%
Total	568	4,954	11%

[a] This table presents descriptive statistics for the 568 sample firm-years with disproportionate inventory increases between 1998 and 2002.

[b] All observations refer to firm-years satisfying sample selection criteria (1)–(3), excluding the 102 firm-years that also meet criteria (4)–(5) but have different inventory numbers on Compustat and in Form 10-K or do not have a Form 10-K on the SEC's EDGAR database.

Variable Definitions:

INVDAYS_CH% = firm *i*'s percentage change in days in inventory (*INVDAYS*), computed as (365/(Compustat #41/Compustat #3)), from year *t*–1 to year *t* (*INVDAYS*/lag(*INVDAYS*)−1);
INV_CH% = firm *i*'s percentage change in inventory from year *t*–1 to year *t* (Compustat #3/lag(Compustat #3) − 1);
MV = market value of common equity of firm *i* at the end of year *t* (Compustat #199 * Compustat #25);
LnMV = natural log of the market value of common equity for firm *i* at the end of year *t*;
BM = book-to-market ratio for firm *i* at the end of year *t* (Compustat #60/(Compustat #199 * Compustat #25));
ROA = firm *i*'s return on total assets in year *t* (Compustat #178/Compustat #6);
StdROA = standard deviation of firm *i*'s non-missing return on total assets from year *t*–4 to year *t*,
ROA_CH = firm *i*'s change in return on total assets from year *t*–1 to year *t* (*ROA* − lag(*ROA*));
BHR = firm *i*'s stock returns for year *t* computed as the buy-and-hold stock returns over the 12 months ending four months after the fiscal year-end of year *t*; monthly stock returns are obtained from CRSP;
StdBHR = standard deviation of firm *i*'s non-missing annual buy-and-hold stock returns from year *t*–4 to year *t*;
NUMEST = number of analysts issuing annual earnings forecasts for firms *i* and for year *t* as reported on I/B/E/S;
PERCINT = percentage of firm *i*'s shares held by institutions at the end of year *t* computed as institutional shareholding at year-end as reported on CDA Spectrum divided by shares outstanding at year-end;
SALESGR = firm *i*'s sales growth rate in year *t* (Compustat #12/lag(Compustat #12) − 1);
BETA = firm *i*'s beta estimated using market-model regression of its daily stock returns on CRSP daily value-weighted market returns over the 12 months ending four months after the fiscal year-end of year *t*;
Std_INVDAYS_CH% = standard deviation of non-missing percentage change in days in inventory (*INVDAYS_CH%*) for firm *i* from year *t*–5 to year *t*–1;
INVDAYS_CH%lag1 = firm *i*'s percentage change in days in inventory from year *t*–2 to year *t*–1;
ROAleadi = firm *i*'s arithmetic average return on assets in the subsequent *i* years (*i* = 1,2,3);
ROA_CHleadi = firm *i*'s arithmetic average change in return on assets in the subsequent *i* years (*i* = 1,2,3);
SALESGRleadi = firm *i*'s geometric average sales growth in the subsequent *i* years (*i* = 1,2,3); and
BHRleadi = firm *i*'s geometric average buy-and-hold stock returns in the subsequent *i* years (*i* = 1,2,3).

TABLE 3

Comparison between Sample Firm-Years (n = 568) and Non-Sample Firm-Years (n = 4,386)[a]

Variables	Non-Sample Mean	Sample Mean	Difference in Mean	t-stat	Difference in Median	Z-stat
Inventory Related						
INVDAYS_CH%	−0.0753	0.4562	−0.5315***	−43.80	−0.4095***	−37.30
Std_INVDAYS_CH%	0.3295	0.3678	−0.0383**	−1.98	−0.0352***	−4.26
INVDAYS_CH%lag1	0.0789	−0.0010	0.0800***	4.66	0.0756***	6.95
Selected Characteristics						
MV	1114.02	550.22	563.80 ***	4.87	18.82	1.40
BM	0.8147	0.8072	0.0075	0.21	0.0431	1.29
StdROA	0.0831	0.0986	−0.0155***	−3.47	−0.0164***	−5.97
StdBHR	0.6004	0.6861	−0.0857***	−3.24	−0.0718***	−4.70
NUMEST[b]	5.2169	4.3333	0.8835***	3.77	0.0000*	1.80
PERCINT[c]	0.3596	0.3246	0.0351***	2.95	0.0611***	2.63
Current Performance						
ROA	0.0256	0.0251	0.0004	0.05	0.0082	0.77
ROA_CH	−0.0117	−0.0166	0.0050	0.80	0.0145***	3.04
SALESGR	0.0602	0.0295	0.0306**	2.49	0.0351***	3.60
BHR	0.1059	0.0714	0.0345	0.83	0.0664*	1.96
Future Performance						
ROAlead1	0.0179	−0.0066	0.0245**	2.58	0.0232***	4.18
ROAlead2	0.0219	−0.0006	0.0225**	2.44	0.0207***	4.11
ROAlead3	0.0272	0.0080	0.0193**	2.32	0.0183***	4.02
ROA_CHlead1	−0.0075	−0.0321	0.0246***	3.83	0.0196***	6.22
ROA_CHlead2	−0.0062	−0.0154	0.0093**	2.23	0.0053***	3.24
ROA_CHlead3	−0.0043	−0.0112	0.0070**	2.18	0.0040***	3.10
SALESGRlead1	0.0367	0.0983	−0.0616***	−4.09	−0.0319***	−3.99
SALESGRlead2	0.0382	0.0678	−0.0296***	−2.78	−0.0142**	−2.51
SALESGRlead3	0.0398	0.0553	−0.0156	−1.61	−0.0009	−0.53
BHRlead1	0.2447	0.1808	0.0639	1.41	0.0700***	3.16
BHRlead2	0.0935	0.0455	0.0480**	2.25	0.0571***	3.34
BHRlead3	0.0936	0.0437	0.0499***	2.95	0.0549***	3.57

*, **, *** Indicate statistical significance at the 10 percent, the 5 percent, and the 1 percent confidence levels, respectively (two-tailed).

[a] This table compares selected firm characteristics and performance measures between sample firm-years and non-sample firm-years. Non-sample firm-years are those satisfying sample selection criteria (1)–(3), but with an increase in days in inventor of less than 20 percent or without an increase in the inventory level. There are 4,386 such observations. Average firm characteristics for sample firm-years and non-sample firm-years, differences in means, and differences in medians are reported. Test for differences in means are based on t-statistics (two-tailed). Tests for differences in medians are based on Z-statistics (two-tailed) from Wilcoxon rank sum test. Variables are defined in Table 2.

[b] 1,578 out of 4,386 non-sample firm-years and 208 out of 568 sample firm-years are not followed by analysts as reported on I/B/E/S.

[c] 41 out of 4,386 non-sample firm-years and 3 out of 568 sample firm-years do not have shares held by institutions as reported on CDA Spectrum.

firms with disproportionate inventory increases includes not only firms with deteriorating performance, but also those with growth potential. To differentiate between these two types of firms, financial information users demand additional information.

IV. RESEARCH DESIGN AND RESULTS

MD&A Disclosures on Inventory Increases

I read the MD&A for each of the 568 firm-year observations in the sample and collect disclosures on inventory changes. Because assessing the quality of narrative disclosures is subjective, I only code the existence, and the favorability, of MD&A disclosures on inventory increases. The disclosure-existence score, *INVDIS_EXS*, takes a value of 1 if the firm explains inventory increases in MD&A, and 0 otherwise. To control for the possibility that a firm can discuss inventory changes in other part(s) of its Form 10-K, I read all statements on inventory in Form 10-K. I find that most inventory-change disclosures are provided under "liquidity and capital resources" or "results of operations" in MD&A. Among the 568 sample firm-years, only three discuss inventory changes in footnotes to financial statements and three discuss them in business overview. For empirical tests, I treat these six observations as firms explaining inventory increases in MD&A. The results are qualitatively similar if I exclude these six observations.

I divide the subsample of firms that explain inventory increases into three groups depending on the favorability of the provided explanations. The favorability is determined by the direction of the demand shift suggested by the explanations. Appendix A describes the coding system of MD&A disclosures on inventory increases. Favorable explanations include anticipation of high future demand, expectation of sales growth, introduction of new product(s) in the current or the subsequent year, receipt of new contract(s), and addition of new retail store(s). Such statements indicate that sales will increase in the future and, thus, represent good news. If a firm states that the inventory level increases to support sales growth in the current period and if there is no specific indication in Form 10-K that sales will decrease in the subsequent year, then this firm is also classified as having favorable explanations for inventory increases. The favorable-inventory-disclosure variable, *INVDIS_FAV*, takes a value of 1 for firm-years with favorable explanations for disproportionate inventory increases, and 0 otherwise.

Examples of unfavorable disclosures include lower than expected sales, decreases in orders or sales in the current period, cancellation of orders, and products returned from certain customer(s) or distributor(s). These explanations suggest that sales in the current period are lower than the expectation or past sales. A downward demand shift and a significant inventory buildup increase the likelihood of inventory obsolescence and can lead to an inventory write-off that reduces profit. Therefore, these explanations are classified as unfavorable. The unfavorable-inventory-disclosure variable, *INVDIS_UNF*, equals 1 for observations with unfavorable explanations for disproportionate inventory increases, and 0 otherwise.

Neutral explanations for inventory increases include improvement in customer service, placement of more trial and demonstration machines at customer locations, addition of new showrooms, an increase in the amount of safety stock of inventory, a shift in the timing of inventory purchase or production, an increase in inventory to support a higher level of production, a volume purchase of inventory items at a favorable price, a planned buildup in inventory in anticipation of price changes, and a lower than normal inventory level at the end of the previous year. Such explanations are not directly associated with a firm's current expectation for future sales or with the relation between current sales and past expectation (for current sales). Interpretations of these explanations are ambiguous. For example, the decision to improve customer service can be related to several reasons. It is possible that a firm wants to improve customer satisfaction in order to have

a positive sales growth rate in the coming year. However, it is also possible that a firm foresees problems with its products and tries to reduce the sales decrease by improving customer service. Therefore, explanations of this type are classified as neutral.

Table 4 reported descriptive statistics for MD&A disclosures on disproportionate inventory increases. Among the 568 firm-years with disproportionate inventory increases, 282 (49.65 percent) explain the changes in MD&A (Panel A). Thus, a significant number of firms with disproportionate inventory increases choose not to explain inventory changes in MD&A. Among the 282 observations that explain inventory changes, 119 (42.20 percent) have favorable explanations, 105 (37.23 percent) provide neutral explanations, and 58 (20.57 percent) give unfavorable explanations (Panel B). Untabulated analyses present no evidence that observations in a particular sub-industry are more likely to explain inventory changes or are more likely to provide explanations of a certain type (favorability). However, 67 percent of sample firms in 2002 explain inventory increases, while 48 percent, on average, explain the changes in other years. The increase in the percentage of firms that explain inventory increases in 2002 can be a result of several reasons. First, the Sarbanes-Oxley Act, passed on July 30, 2002, requires management to review and certify financial reports since August 29, 2002. Managers may thus have an incentive to provide a more complete discussion in annual reports. Second, the SEC announced in December 2001 that it would review the annual reports filed in 2002 by the Fortune 500 companies to identify inappropriate or deficient disclosures. The SEC also issued a statement about its view regarding the purpose and the content of MD&A disclosures in January 2002. The SEC's attempt to discipline MD&A disclosures can motivate more firms to explain disproportionate inventory increases in MD&A.

Forward-Looking Statements in Form 10-K on a Firm's Performance Outlook

Prior research (e.g., Bryan 1997; Cole and Jones 2004) finds that MD&A disclosures contain information associated with future firm performance. To control for the possibility that the information in MD&A inventory disclosures is correlated with general forward-looking information in MD&A or in Form 10-K, I identify forward-looking statements in Form 10-K on a firm's performance outlook and assess the overall favorability of these statements for each observation in the sample.[8] A firm can provide forward-looking statements on many aspects of its operations, such as sales, costs, strategy, and competition. To assess the overall favorability of a firm's performance outlook, I focus on future profitability or the change in future profitability implied by the forward-looking statements. If these statements suggest that a firm's profit or profit margin will be maintained at a similar level or even increase, the favorable-outlook variable, *FUTURE_FAV*, takes a value of 1. If a firm indicates that its profitability will decrease or if negative income is expected, the unfavorable-outlook variable, *FUTURE_UNF*, equals 1. If the forward-looking statements do not have consistent implications for a firm's future profitability or if a firm does not have specific predictions, then the firm's outlook is expected to be neutral. Appendix B details the classification of forward-looking statements on a firm's performance outlook. The forward-looking statements are favorable for 146 (25.70 percent) of the 568 firm-years in the sample, neutral for 297 (52.29 percent) observations, and unfavorable for 125 (22.01 percent) firm-years (Table 4, Panel C).

Table 4, Panel D presents a contingency table of the favorability of forward-looking statements in Form 10-K and the type of inventory disclosures in MD&A. I find that the favorability of a firm's forward-looking statements and the favorability of its explanations for the disproportionate inventory increase are correlated. For example, 49 percent of firm-years with favorable inventory disclosures also have a favorable performance outlook. However, at the same time, they

[8] The forward-looking statements in 10-K are identified by searching for sentences containing "believe," "anticipate," "expect," "estimate," "plan," "seek," "intend," "project," "outlook," or "goal."

TABLE 4

Descriptive Statistics for Disclosures on Inventory Increases in MD&A and Forward-Looking Statements in Form 10-K[a]

Panel A: Existence of MD&A Disclosures on Inventory Increases

	Absent	Existent	Total
Number of Observations	286	282	568
Percentage	50.35%	49.65%	100%

Panel B: Favorability of MD&A Disclosures on Inventory Increases

	Unfavorable	Neutral	Favorable	Total
Number of Observations	58	105	119	282
Percentage	20.57%	37.23%	42.20%	100%

Panel C: Favorability of Forward-Looking Statements in Form 10-K on a Firm's Performance Outlook

	Unfavorable	Neutral	Favorable	Total
Number of Observations	125	297	146	568
Percentage	22.01%	52.29%	25.70%	100%

Panel D: Contingency Table of the Favorability of Forward-Looking Statements in Form 10-K and the Type of MD&A Disclosures on Inventory Increases

		MD&A Inventory Disclosures							
		Unfavorable		Neutral		Favorable		Absent	
		n	%	n	%	n	%	n	%
Forward-Looking Statements	Unfavorable	30	52%	12	11%	12	10%	71	25%
	Neutral	25	43%	64	61%	49	41%	159	56%
	Favorable	3	5%	29	28%	58	49%	56	19%
		58	100%	105	100%	119	100%	286	100%

[a] This table presents descriptive statistics for MD&A disclosures on inventory increases and forward-looking statements in Form 10-K on a firm's performance outlook. Please refer to Appendix A for examples of MD&A disclosures on inventory increases and Appendix B for examples of forward-looking statements in Form 10-K on a firm's performance outlook.

are likely to contain information incremental to each other. Among the 119 firm-years with favorable inventory disclosures, 41 percent (10 percent) have a neutral (unfavorable) outlook because expected increases in costs offset or more than offset expected increases in sales.

Tests of Hypotheses

Predictive Ability of MD&A Inventory Disclosures for Future Firm Performance (H1)

To investigate the predictive ability of MD&A explanations for disproportionate inventory increases for future firm performance (H1), I examine the association between inventory-disclosure variables and average financial performance measures—return on assets (*ROA*) and sales growth (*SALESGR*)—in the subsequent one, two, or three years using Regressions (1) and (2).

$$\begin{aligned} ROAleadi_t = {} & a_i + b_{i1}INVDIS\_EXS + b_{i2}INVDIS\_FAV + b_{i3}INVDIS\_UNF + b_{i4}FUTURE\_FAV \\ & + b_{i5}FUTURE\_UNF + b_{i6}INVDAYS\_CH\% + b_{i7}INV\_CH\% + b_{i8}SALESGR \\ & + b_{i9}ROA + b_{i10}ROA\_CH + b_{i11}StdROA + b_{i12}LnMV + b_{i13}BM + c_iYearIndicators \\ & + e_i(i = 1,2,3); \end{aligned} \tag{1}$$

$$\begin{aligned} SALESGRleadi_t = {} & a_i + b_{i1}INVDIS\_EXS + b_{i2}INVDIS\_FAV + b_{i3}INVDIS\_UNF \\ & + b_{i4}FUTURE\_FAV + b_{i5}FUTURE\_UNF + b_{i6}INVDAYS\_CH\% \\ & + b_{i7}INV\_CH\% + b_{i8}SALESGR + b_{i9}TA\_CH\% + b_{i10}ROA + b_{i11}ROA\_CH \\ & + c_iYearIndicators + e_i(i = 1,2,3). \end{aligned} \tag{2}$$

I examine ROA because profitability is the focus of financial statement analysis. I examine sales growth because sales are a necessary condition for operating profits for manufacturing firms and because the favorability of MD&A inventory disclosures is defined based on the direction of demand shifts suggested by the explanations.

The variables of interest in Regressions (1) and (2) are the variables related to inventory disclosures in MD&A. *INVDIS_EXS* indicates the existence of MD&A explanations for inventory increases. It takes a value of 1 if a firm explains the inventory changes, and 0 otherwise. Its coefficient b_{i1} represents the performance difference between firms that provide neutral explanations and those that do not explain the inventory changes. *INVDIS_FAV* takes a value of 1 if a firm explains the inventory increase in MD&A and if the provided explanations are favorable, and 0 otherwise. Similarly, *INVDIS_UNF* equals 1 if a firm provides unfavorable explanations for its inventory increase, and 0 otherwise. The coefficient b_{i2} (b_{i3}) on *INVDIS_FAV* (*INVDIS_UNF*) measures the performance difference between firms with favorable (unfavorable) explanations and those with neutral explanations. To test H1, I focus on the coefficient b_{i2} on *INVDIS_FAV* and the coefficient b_{i3} on *INVDIS_UNF*. The null hypothesis that the favorability of MD&A explanations is not associated with future firm performance implies $b_{i2} = b_{i3} = 0$. Evidence of $b_{i2} > 0$, $b_{i3} < 0$, and/or $b_{i2} > b_{i3}$ would suggest that future firm performance is positively associated with the favorability of MD&A inventory disclosures.[9] Although I do not examine the implications of a firm's decision to explain the inventory increase in MD&A for its prospects (H2) with Regressions (1) and (2), the coefficients on inventory disclosure variables in these regressions imply the

[9] Treating the MD&A inventory-disclosure variables as endogenous is inappropriate because instead of examining whether the disclosures have any impact on future firm performance, I investigate whether firms that explain the inventory changes or provide a certain type of explanations differ in terms of future financial performance. I examine the association between disclosures and future performance, instead of a causal relation between them.

prospects of non-disclosure firms, as compared to those of firms with favorable or unfavorable disclosures. In particular, $b_{i1} + b_{i2}$ ($b_{i1} + b_{i3}$) measures the performance difference between firms with favorable (unfavorable) explanations and non-disclosure firms.

Following Core et al. (1999) and Hanlon et al. (2003), in Regression (1), average ROA in the subsequent i years (*ROAleadi*) is regressed on the inventory disclosure variables, ROA in the current year (*ROA*), and control variables. Two indicator variables related to statements in Form 10-K on a firm's performance outlook are also included to control for general forward-looking information in Form 10-K. *FUTURE_FAV* (*FUTURE_UNF*) takes a value of 1 if a firm's forward-looking statements in Form 10-K are favorable (unfavorable), and 0 otherwise. To examine whether MD&A disclosures on inventory increases augment the information on inventory changes available in financial statements, I control for the percentage change in days in inventory (*INV-DAYS_CH%*) and the percentage change in inventory (*INV_CH%*). To address the possibility that the favorability of MD&A inventory disclosures is correlated with current sales growth, I control for sales growth in year t (*SALESGR*). I use the level of ROA and the change in ROA (*ROA_CH*) in year t as additional independent variables to control for their predictive ability for future ROA. To control for the relation between firm risk and future profitability, I also include the standard deviation of ROA from year t–4 to year t (*StdROA*), the natural logarithm of market value of common equity at the end of year t (*LnMV*), and the book-to-market ratio at the end of year t (*BM*). I use year indicators to control for the time variation in ROA.

In Regression (2) for future sales growth (*SALESGRleadi*), besides the disclosure variables, the two inventory-change variables (*INV_CH%, INVDAYS_CH%*), and sales growth in year t (*SALESGR*), I include the percentage change in total assets from year t–1 to year t (*TA_CH%*), ROA in year t (*ROA*), and the change in ROA in year t (*ROA_CH*) as another three independent variables. *TA_CH%* controls for the positive relation between the current change in total assets and future sales growth. As the firm grows, sales are likely to increase. *ROA* and *ROA_CH* control for the effect of current profitability or change in profitability on future competition. A high profitability level or an increase in profitability attracts competition and thus may lead to a lower sales growth rate in the future.

Estimation results for Regression (1) are reported in Table 5, Panel A. The coefficient on the indicator of favorable MD&A inventory disclosures (*INVDIS_FAV*) is positive in all three columns and significant in the regressions for average ROA in the subsequent two or three years (*ROAlead2,3*). The coefficient on the indicator of unfavorable inventory disclosures (*INVDIS-_UNF*) is negative in all three columns and significant in the regression for ROA in year t+1 (*ROAlead1*). These coefficients suggest that firms with favorable (unfavorable) explanations for disproportionate inventory increases have significantly higher (lower) future ROA than firms with neutral explanations. To examine whether firms with favorable explanations have significantly higher future ROA than those with unfavorable explanations, I compare the coefficients on *INVDIS_FAV* and *INVDIS_UNF* (Test 1 (null): *INVDIS_FAV − INVDIS_UNF = 0*). I find that firms with favorable inventory disclosures have significantly higher future ROA in all three columns. For example, *INVDIS_FAV − INVDIS_UNF* equals 0.0790 ($p = 0.001$) in the regression for ROA in year t+1 (*ROAlead1*). 0.0790 represents 58 percent of the interquartile range of *ROAlead1* for all manufacturing firm-years satisfying example selection criteria (1)–(3), suggesting that the performance difference between firms with favorable inventory disclosures and those with unfavorable inventory disclosures is economically significant. These results suggest that MD&A explanations for disproportionate inventory increases, if provided, have predictive ability for ROA in

TABLE 5

Association between Favorability of MD&A Disclosures on Inventory Increases and Future Firm Performance Tests of Hypothesis 1

Panel A: Profitability (*ROA*)[a]

	Dependent Variables		
Independent Variables	***ROAlead1***	***ROAlead2***	***ROAlead3***
Intercept	0.0140	−0.0050	−0.0160
	(0.46)	(−0.16)	(−0.55)
INVDIS_EXS	−0.0020	−0.0110	−0.0080
	(−0.13)	(−0.71)	(−0.49)
INVDIS_FAV	0.0210	0.0350*	0.0330*
	(1.10)	(1.86)	(1.78)
INVDIS_UNF	−0.0580**	−0.0350	−0.0200
	(−2.43)	(−1.47)	(−0.86)
FUTURE_FAV	0.0510***	0.0510***	0.0500***
	(3.48)	(3.59)	(3.53)
FUTURE_UNF	−0.0640***	−0.0500***	−0.0310**
	(−4.00)	(−3.21)	(−2.07)
INVDAYS_CH%	0.0090	−0.0050	0.0030
	(0.33)	(−0.20)	(0.10)
INV_CH%	−0.0460	−0.0360	−0.0240
	(−1.53)	(−1.26)	(−0.83)
SALESGR	0.0550	0.0330	0.0110
	(1.55)	(0.98)	(0.32)
ROA	0.6030***	0.5300***	0.4670***
	(14.75)	(13.28)	(11.72)
ROA_CH	−0.0150	0.0120	−0.0060
	(−0.31)	(0.25)	(−0.13)
StdROA	−0.3030***	−0.2430***	−0.2350***
	(−3.93)	(−3.25)	(−3.05)
LnMV	0.0080*	0.0100**	0.0080**
	(1.85)	(2.50)	(1.95)

(continued on next page)

Panel A: Profitability (*ROA*)[a]

Independent Variables	Dependent Variables		
	ROAlead1	*ROAlead2*	*ROAlead3*
BM	−0.0100	0.0050	0.0040
	(−1.05)	(0.57)	(0.43)
n	568	521	463
Adjusted R^2	0.576	0.552	0.490
Tests on Coefficients			
Test 1: *INVDIS_FAV* − *INVDIS_UNF* > 0			
INVDIS_FAV − *INVDIS_UNF*	0.0790	0.0700	0.0530
Wald-statistic	10.94	8.64	4.97
p-value	0.0010	0.0034	0.0259
Test 2: *INVDIS_FAV* + *INVDIS_EXS* > 0			
INVDIS_FAV + *INVDIS_EXS*	0.0190	0.0240	0.0250
Wald-statistic	1.44	2.37	2.79
p-value	0.2306	0.1234	0.0952
Test 3: *INVDIS_UNF* + *INVDIS_EXS* < 0			
INVDIS_UNF + *INVDIS_EXS*	−0.0600	−0.0460	−0.0280
Wald-statistic	8.18	4.84	1.74
p-value	0.0042	0.0278	0.1876

Panel B: Sales Growth (*SALESGR*)[b]

Independent Variables	Dependent Variables		
	SALESGRlead1	*SALESGRlead2*	*SALESGRlead3*
Intercept	0.0760**	0.0710***	0.0220
	(2.01)	(2.60)	(0.88)
INVDIS_EXS	−0.0400	−0.0360	−0.0290
	(−1.12)	(−1.41)	(−1.23)
INVDIS_FAV	0.1250***	0.0930***	0.0440
	(2.95)	(3.08)	(1.59)

(continued on next page)

Panel B: Sales Growth (*SALESGR*)[b]

Independent Variables	Dependent Variables		
	SALESGRlead1	***SALESGRlead2***	***SALESGRlead3***
INVDIS_UNF	−0.0340	−0.0250	−0.0200
	(−0.63)	(−0.64)	(−0.55)
FUTURE_FAV	0.1340***	0.1010***	0.0700***
	(4.13)	(4.33)	(3.24)
FUTURE_UNF	−0.1140***	−0.0370	−0.0330
	(−3.22)	(−1.43)	(−1.40)
INVDAYS_CH%	−0.0200	−0.0050	0.0480
	(−0.34)	(−0.12)	(1.18)
INV_CH%	0.0720	0.0600	0.0090
	(1.03)	(1.22)	(0.19)
SALESGR	0.0760	−0.0260	0.0710
	(0.96)	(−0.46)	(1.25)
TA_CH%	0.1110**	0.0800**	0.0740**
	(2.37)	(2.34)	(2.29)
ROA	−0.4340***	−0.2240***	−0.1740***
	(−5.78)	(−4.18)	(−3.36)
ROA_CH	0.0890	0.0410	−0.0430
	(0.83)	(0.54)	(−0.58)
n	568	521	462
Adjusted R^2	0.235	0.166	0.178
Tests on Coefficients			
Test 1: *INVDIS_FAV − INVDIS_UNF* > 0			
INVDIS_FAV − INVDIS_UNF	0.1590	0.1180	0.0640
Wald-statistic	8.78	9.07	3.05
p-value	0.0030	0.0026	0.0810
Test 2: *INVDIS_FAV + INVDIS_EXS* > 0			
INVDIS_FAV + INVDIS_EXS	0.0850	0.0570	0.0150
Wald-statistic	5.82	5.15	0.42
p-value	0.0158	0.0234	0.5184
Test 3: *INVDIS_UNF + INVDIS_EXS* < 0			

(continued on next page)

Panel B: Sales Growth (*SALESGR*)[b]

	Dependent Variables		
Independent Variables	***SALESGRlead1***	***SALESGRlead2***	***SALESGRlead3***
INVDIS_UNF + INVDIS_EXS	−0.0740	−0.0610	−0.0490
Wald-statistic	2.49	3.13	2.28
p-value	0.1148	0.0768	0.1308

*, **, *** Indicate statistical significance at the 10 percent, the 5 percent, and the 1 percent confidence levels, respectively (two-tailed).

[a] Panel A presents the coefficient estimates from OLS regressions of a firm's future profitability (average *ROA* in the subsequent one, two, or three years) on MD&A-inventory-disclosure variables and control variables. The sample consists of 568 firm-years with disproportionate inventory increases during 1998–2002. Among these 568 firm-years, 282 observations explain the inventory changes in MD&A. In particular, 58 provide unfavorable explanations, 105 neutral, and 119 favorable. Indicator variables for year 1999, 2000, 2001, and 2002 are included as control variables. Coefficients and t-statistics (in parentheses) are reported. Tests of the hypothesized relation between coefficients on MD&A-inventory-disclosure variables are reported below the regression estimates.

[b] Panel B presents the coefficient estimates from OLS regressions of a firm's future sales growth (average sales growth in the subsequent one, two, or three years) on MD&A-inventory-disclosure variables and control variables. The sample consists of 568 firm-years with disproportionate inventory increases during 1998–2002. Among these 568 firm-years, 282 observations explain the inventory changes in MD&A. In particular, 58 provide unfavorable explanations, 105 neutral, and 119 favorable. Indicator variables for year 1999, 2000, 2001, and 2002 are included as control variables. Coefficients and t-statistics (in parentheses) are reported. Tests of the hypothesized relation between coefficients on MD&A-inventory-disclosure variables are reported below the regression estimates.

Variable Definitions:
INVDIS_EXS = 1 if firm *i* explains inventory increases in MD&A in year *t*, and 0 otherwise;
INVDIS_FAV = 1 if firm *i*'s MD&A disclosures on inventory increases are favorable in year *t*, and 0 otherwise;
INVDIS_UNF = 1 if firm *i*'s MD&A disclosures on inventory increases are unfavorable in year *t*, and 0 otherwise;
FUTURE_FAV = 1 if firm *i*'s forward-looking statements on its performance outlook in Form 10-K are favorable in year *t*, and 0 otherwise; and
FUTURE_UNF= 1 if firm *i*'s forward-looking statements on its performance outlook in Form 10-K are unfavorable in year *t*, and 0 otherwise.
Other variables are defined in Table 2.

the subsequent three years, rejecting null hypothesis of H1.[10]

For control variables, the coefficient on the indicator of favorable forward-looking information in Form 10-K (*FUTURE_FAV*) is significantly positive and the coefficient on the indicator of unfavorable forward-looking statements in Form 10-K (*FUTURE_UNF*) is significantly negative, confirming the results in prior studies (e.g., Bryan 1997) on the predictive ability of narrative disclosures in MD&A by using data in a more recent period, although *FUTURE_FAV* and *FUTURE_UNF* are defined using forward-looking statements in Form 10-K. One difference in research design between Bryan (1997) and this study is that Bryan (1997) uses the direction of change in future sales or earnings per share (*EPS*) as the dependent variable, while this study uses the level of future ROA or sales growth directly. The later approach enables an examination of the magnitude of the association between current disclosures and future performance. For example, results in Table 5, Panel A indicate that average ROA in year $t+1$ (*ROAlead1*) of firms with favorable (unfavorable) forward-looking disclosures in Form 10-K is higher by 0.0510 (lower by 0.0640) than firms with neutral disclosures. These performance differences are economically significant because they represent 37 percent and 47 percent of the interquartile range of *ROAlead1* for all manufacturing firm-years satisfying example selection criteria (1)–(3).[11]

The coefficients on the percentage change of days in inventory or inventory (*INVDAYS_CH%* and *INV_CH%*) are not significant in any column, suggesting that, for firms with disproportionate inventory increases, inventory changes have varying implications for future profitability. The coefficient on *ROA* is significant and between 0 and 1, consistent with mean reversion of accounting performance measures. I also find that firms with a greater variation in ROA over time (*StdROA*) have a lower future ROA, while firms with a larger size (*LnMV*) have a higher future ROA.

Table 5, Panel B reports the estimation results for Regression (2). For H1, the coefficient on the indicator of favorable inventory disclosures (*INVDIS_FAV*) is significantly positive in the regressions for average sales growth in the subsequent one or two years (*SALESGRlead1,2*); the coefficient on the indicator of unfavorable inventory disclosures (*INVDIS_UNF*) is negative but insignificant in all three columns. The test on the performance difference between firms with favorable explanations and those with unfavorable explanations (Test 1) suggests that favorable-disclosure firms have significantly higher sales growth in subsequent years. For example, *INVDIS_FAV* − *INVDIS_UNF* is 0.1590 in the regression for *SALESGRlead1*, and 0.1590 is 69 percent of the interquartile range of *SALESGRlead1* for all manufacturing firm-years satisfying sample selection criteria (1)–(3). These results suggest that MD&A explanations for disproportionate inventory increases have predictive ability for a firm's future sales growth, again rejecting the null of H1.

The coefficient on the indicator of favorable forward-looking statements (*FUTURE_FAV*) is significantly positive in all three columns, and the coefficient on the indicator of unfavorable

[10] Bernard and Noel (1991) show that changes in different inventory components have different implications for future earnings and sales. In particular, unexpected changes in raw materials and work in process are positive predictors for future sales, while unexpected changes in finished goods are negative predictors for future earnings for manufacturers. To examine whether MD&A inventory disclosures have predictive ability for future ROA and sales growth after controlling for changes in inventory components, I use the 381 firm-years that disclose inventory components in both the current year and the previous year and replace inventory change variables in Regressions (1) and (2) by changes in inventory components. The results for hypotheses are qualitatively similar.

[11] Although results in Table 5 and statistics in Table 4, Panel C both suggest that inventory disclosures in MD&A and forward-looking statements in 10-K contain incremental information, I conduct additional analysis to examine whether MD&A inventory disclosures unrelated to predictions for the future have predictive ability for firm prospects. After removing 32 observations that attribute inventory increases to forecasted future sales change, I still find that firms with favorable MD&A inventory disclosures have better accounting outcomes in the subsequent years than those with unfavorable disclosures, suggesting that MD&A inventory disclosures related to past events are informative.

forward-looking information (*FUTURE_UNF*) is significantly negative in the regression for *SALESGRlead1*, suggesting that forward-looking statements, though generally qualitative, contain information associated with future sales growth. The two inventory-change variables, *INVDAYS_CH%* and *INV_CH%*, do not have a significant coefficient in any column, suggesting that disproportionate inventory increases do not have uniform implications for future sales growth. Consistent with my expectation, the coefficient on the change in total assets (*TA_CH%*) is positive and the coefficient on *ROA* is negative, indicating that sales increase as the firm size grows and decrease as competition intensifies.

In summary, results in Table 5 show that the favorability of MD&A disclosures on inventory increases has predictive ability for a firm's profitability and sales growth in the subsequent three years, implying that MD&A explanations for inventory increases are reliable on average and contain relevant information for financial statement analysis. Untabulated analyses show that 97 (82 percent) out of the 119 firm-years with favorable explanations have positive sales growth in the subsequent year. Among the 105 and the 58 firm-years with neutral or unfavorable explanations, only 58 (55 percent) and 15 (26 percent) have positive sales growth in the subsequent year, respectively. This evidence provides further support that MD&A disclosures on inventory increases are not cheap talk.

Prior research finds that disproportionate inventory changes are negatively associated with a firm's current and future stock returns and future earnings (Lev and Thiagarajan 1993; Abarbanell and Bushee 1997, 1998). However, the results from testing H1 suggest that the disproportionate inventory increase is not necessarily a negative signal for a specific firm. The group of firms with disproportionate inventory increases consists of firms with good prospects and firms with poor prospects. For example, firms with favorable (unfavorable) MD&A inventory disclosures have an average ROA of 5.90 percent (−11.61 percent) in the following year, well above (below) the industry average of 1.51 percent. The H1 results imply that investors can potentially differentiate between these two types of firms and make better financial statement analyses by looking at relevant disclosures in MD&A. Such results reinforce the recommendation by Penman (2001) to examine firm disclosures to better understand changes in accounting ratios.[12]

Implications of the Decision to Explain Inventory Increases for Future Performance (H2)

I investigate the implications of a firm's decision to explain its disproportionate inventory increase in MD&A for its future performance (H2) by examining the association between the existence of MD&A inventory disclosures and average return on assets (*ROA*) or sales growth (*SALESGR*) in the subsequent one, two, or three years using Regressions (3) and (4).

$$\begin{aligned} ROAleadi_t = {} & a_i + b_{i1}INVDIS\_EXS + b_{i2}FUTURE\_FAV + b_{i3}FUTURE\_UNF \\ & + b_{i4}INVDAYS\_CH\% + b_{i5}INV\_CH\% + b_{i6}SALESGR + b_{i7}ROA + b_{i8}ROA\_CH \\ & + b_{i9}StdROA + b_{i10}LnMV + b_{i11}BM + c_i YearIndicators + e_i (i = 1,2,3). \end{aligned} \quad (3)$$

[12] To better understand the link between current inventory changes and future earnings and sales growth, in untabulated analyses, I compare several other performance outcomes between firms with different types of MD&A inventory disclosures. I find that firms with favorable explanations have an increase in gross margin in year *t*, an increase in labor force starting in year *t* and a smaller decrease in days in inventory in the following years. These results indicate that because current inventory increases for firms with favorable explanations are related to good reasons (e.g., expectation of sales growth), these firms can charge a higher sales price, hire more employees, and maintain an inventory level suggested by its performance expectation. These results also indirectly suggest that the classification of MD&A inventory disclosures in this study is reasonable.

$$SALESGRlead i_t = a_i + b_{i1}INVDIS\_EXS + b_{i2}FUTURE\_FAV + b_{i3}FUTURE\_UNF + b_{i4}INVDAYS\_CH\% + b_{i5}INV\_CH\% + b_{i6}SALESGR + b_{i7}TA\_CH\% + b_{i8}ROA + b_{i9}ROA\_CH + c_i YearIndicators + e_i (i = 1,2,3). \quad (4)$$

Regressions (3) and (4) are similar to Regressions (1) and (2), except that the existence of MD&A inventory disclosures (*INVDIS_EXS*) is the only MD&A-inventory-disclosure variable included. Evidence of a significant coefficient b_{i1} on *INVDIS_EXS* would reject the null of H2 (i.e., $b_{i1} = 0$). A significantly positive coefficient would suggest that, on average, firms that explain inventory increases in MD&A have better future performance than those that do not, consistent with the conjecture that non-disclosure firms are more likely to be earnings manipulators or firms with negative business development.

The estimated coefficients for disclosure variables in Regressions (3) and (4) are presented in Table 6, Panel A and Panel B, respectively. The results for other variables are similar to those in Table 5 and are omitted from Table 6. For H2, the coefficient on the existence of MD&A inventory disclosures (*INVDIS_EXS*) is not statistically significant in any column of either panel, indicating that future profitability or sales growth of firms that explain disproportionate inventory increases in MD&A is not significantly different from that of firms that do not explain the changes. Thus, a firm's decision not to explain disproportionate inventory increases in MD&A does not necessarily indicate bad news. This result suggests that a significant number of firms with favorable private information may choose to pool with firms with low-quality earnings or bad news and choose not to explain inventory increases.[13]

Although non-disclosure firms, as a group, do not have significantly different future performance from disclosure firms, they may have different prospects from firms with favorable or unfavorable explanations. The coefficients on MD&A disclosure variables in Regressions (1) and (2) provide a way to test this conjecture statistically. The sum of the coefficients on the indicator of favorable (unfavorable) inventory disclosures, *INVDIS_FAV* (*INVDIS_UNF*), and on the existence of inventory disclosures, *INVDIS_EXS*, represents the performance difference between firms with favorable (unfavorable) explanations and non-disclosure firms. In Table 5, Panel A, the sum of the coefficients on *INVDIS_FAV* and *INVDIS_EXS* is significantly positive in the regression for average ROA in the subsequent three years, *ROAlead3* (Test 2), and the sum of the coefficients on *INVDIS_UNF* and *INVDIS_EXS* is significantly negative for *ROAlead1* and *ROAlead2* (Test 3). For example, *INVDIS_FAV + INVDIS_EXS* is 0.0190 in the regression for *ROAlead1*, suggesting that average ROA in year t+1 for firms with favorable inventory disclosures is higher than that for non-disclosure firms by 0.0190. In Table 5, Panel B, the sum of the coefficients on *INVDIS_FAV* and *INVDIS_EXS* is significantly positive in the regressions for average sales growth in the following one or two years, *SALESGRlead1* and *SALESGRlead2* (Test 2), and the sum of the coefficients on *INVDIS_UNF* and *INVDIS_EXS* is significantly negative for *SALESGRlead2* (Test 3). These results suggest that, although a firm's decision not to explain the disproportionate inventory increase in MD&A does not necessarily indicate bad news, future financial performance of non-disclosure firms falls between that of firms with favorable explanations and firms with unfavorable explanations.

[13] The insignificant coefficient on the existence of MD&A inventory disclosures (*INVDIS_EXS*) can arise because the group of non-disclosure firms contain both firms with bad news and firms with good news, as discussed in H2. However, such an insignificant result is also consistent with the possibility that non-disclosure firms have little news. Under this conjecture, non-disclosure firms are expected to experience smaller changes in future performance than disclosure firms. However, the data suggest that non-disclosure firms and disclosure firms have similar means or standard deviations for absolute change in ROA or sales growth in the subsequent years.

TABLE 6

Association between Existence of MD&A Disclosures on Inventory Increases and Future Firm Performance Tests of Hypothesis 2

Panel A: Profitability (*ROA*)[a]

	Dependent Variables		
Independent Variables	***ROAlead1***	***ROAlead2***	***ROAlead3***
INVDIS_EXS	−0.0056	−0.0033	0.0030
	(−0.47)	(−0.28)	(0.25)
FUTURE_FAV	0.0586***	0.0594***	0.0574***
	(4.07)	(4.23)	(4.19)
FUTURE_UNF	−0.0745***	−0.0577***	−0.0353**
	(−4.77)	(−3.73)	(−2.35)
n	568	521	463
Adjusted R^2	0.569	0.546	0.486

Panel B: Sales Growth (*SALESGR*)

	Dependent Variables		
Independent Variables	***SALESGRlead1***	***SALESGRlead2***	***SALESGRlead3***
INVDIS_EXS	0.0041	−0.0014	−0.0140
	(0.15)	(−0.07)	(−0.78)
FUTURE_FAV	0.1576***	0.1171***	0.0802***
	(4.90)	(5.09)	(3.79)
FUTURE_UNF	−0.1280***	−0.0458*	−0.0375
	(−3.68)	(−1.81)	(−1.62)
n	568	521	462
Adjusted R^2	0.200	0.148	0.146

*, **, *** Indicate statistical significance at the 10 percent, the 5 percent, and the 1 percent confidence levels, respectively (two-tailed).

[a] This table presents the coefficients on disclosure variables estimated from OLS regressions of a firm's future performance measures on existence of MD&A inventory disclosures

(continued on next page)

and control variables. The dependent variable is average *ROA* (Panel A) and average sales growth (Panel B) in the subsequent one, two, or three years. The estimated coefficients on other variables are similar to those in Table 5 and omitted here. The sample consists of 568 firm-years with disproportionate inventory increases during 1998–2002. Among these 568 firm-years, 282 observations explain the inventory changes in MD&A. Coefficients and t-statistics (in parentheses) are reported.

Variable Definitions:
INVDIS_EXS = 1 if firm *i* explains inventory increases in MD&A in year *t*, and 0 otherwise;
FUTURE_FAV = 1 if firm *i*'s forward-looking statements on its performance outlook in Form 10-K are favorable in year *t*, and 0 otherwise; and
FUTURE_UNF= 1 if firm *i*'s forward-looking statements on its performance outlook in Form 10-K are unfavorable in year *t*, and 0 otherwise.

Determinants of the Decision to Explain Disproportionate Inventory Increases in MD&A

This study's findings imply that MD&A inventory disclosures help users understand the implications of disproportionate inventory increases for future firm performance. However, not all firms with disproportionate inventory increases choose to explain such inventory changes in MD&A. A comprehensive investigation of the determinants of a firm's disclosure decision is beyond the scope of this study. Preliminary tests suggest that a firm's disclosure decision is positively associated with the percentage change in inventory and product differentiation, while negatively associated with performance variability and sales growth. I also find that the propensity to explain inventory changes in MD&A is greater in 2002, probably because of the Sarbanes-Oxley Act passed in 2002 and the SEC's effort in 2001 and 2002 to discipline MD&A disclosures. These results suggest that a firm's disclosure decision is made after considering both disclosure benefits and disclosure costs and can potentially explain why financial performance of non-disclosers is no worse than that of disclosers and why the existence of MD&A inventory disclosures is not necessarily an indicator of better future firm performance.

V. CONCLUSION

This study examines the association between the existence, and the favorability, of MD&A disclosures on disproportionate inventory increases and a firm's future financial performance. Disproportionate inventory increases can arise for a variety of reasons. To understand the inventory changes and to evaluate firm performance, investors can utilize additional information from the firm. This study investigates whether MD&A inventory disclosures can help users interpret disproportionate inventory increases.

I find that the favorability of explanations for disproportionate inventory increases in MD&A is positively associated with a firm's profitability and sales growth in the subsequent three years. However, 286 (50.35 percent) of the 568 manufacturing firm-years with disproportionate inventory increases choose not to explain the changes in MD&A. I find that, although non-disclosure firms do not have significantly poorer future financial performance than disclosure firms, non-disclosure firms' future profitability and sales growth fall between those of favorable-disclosure firms and unfavorable-disclosure firms. These results suggest that the existence, and the favorability, of MD&A inventory disclosures have predictive ability for future financial performance for firms with disproportionate inventory increases.

By examining the association between MD&A disclosures on changes in a line item directly reported on the balance sheet and a firm's future financial performance, this study contributes to the literature on disclosures and shows that MD&A disclosures on inventory increases help users interpret the inventory changes and have predictive ability for future firm performance. It adds to the literature on inventory by showing that disproportionate inventory increases can have different implications for a firm's prospects. This study also contributes to the literature on financial statement analysis by showing that after disproportionate inventory increases are identified, financial information users can use MD&A disclosures to further identify the reasons and the implications.

The Sarbanes-Oxley Act was passed in 2002 to improve the quality of financial disclosures. As this study examines MD&A disclosures on inventory increases during 1998–2002, future study could examine MD&A inventory disclosures in recent years to provide further evidence on whether firms' propensity to explain disproportionate inventory increases in MD&A is greater after the passage of the Sarbanes-Oxley Act and whether these disclosures do a better job in helping users interpret inventory changes. Future studies could also examine MD&A disclosures on other financial statement accounts (e.g., accounts receivable, property, plant, and equipment [PPE]) to gain more insights into the incentives behind MD&A disclosures and the usefulness of these disclosures.

APPENDIX A
The Coding System of MD&A Disclosures on Inventory Increases

I. The Existence of MD&A Disclosures on Inventory Increases

	Description
Existent	There are narrative explanations for inventory increases in MD&A.
Absent	There is no narrative explanation for inventory increases in MD&A.

II. The Favorability of MD&A Disclosures on Inventory Increases for the Subsample of Firms that Explain the Inventory Changes

	Description
Favorable	Anticipation of high future demand, expectation of sales growth, new product introduction in the current or the subsequent year, receipt of new contract(s), addition of new retail store(s), sales growth in the current period. Combination of favorable and neutral statements.
Neutral	Improvement in customer service, placement of more trial and demonstration machines at customer locations, addition of new showrooms, an increase in the amount of safety stock of inventory, a shift in the timing of inventory purchase or production, an increase in inventory to support a higher level of production, a volume purchase of inventory items at a favorable price, a planned buildup in inventory in anticipation of price changes, lower than normal inventory level at the end of the previous year.
Unfavorable	Lower than expected sales, a decrease in orders, a decrease in sales, cancellation of orders, products returned from certain customer(s) or distributor(s). Combination of unfavorable and neutral statements.

APPENDIX B
The Classification of Forward-Looking Statements in Form 10-K on a Firm's Performance Outlook

	Description
Favorable	An expected increase in demand/product price/sales/margins, an expected decrease in cost/expense, new product introduction, an expected increase in profitability, expected positive consequences of litigation resolution or regulation change, etc.
Neutral	No specific prediction or a combination of favorable and unfavorable statements.
Unfavorable	An expected decrease in demand/product price/sales/ margins, an expected increase in cost/expense, an expected decrease in profitability, expected negative consequences of litigation resolution or regulation change, etc.

REFERENCES

Abarbanell, J., and B. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* 35 (1): 1–24.

———, and ———. 1998. Abnormal returns to a fundamental analysis strategy. *The Accounting Review* 73 (1): 19–46.

Barron, O., C. Kile, and T. O'Keefe. 1999. MD&A quality as measured by the SEC and analysts' earnings forecasts. *Contemporary Accounting Research* 16 (1): 75–109.

Bernard, V., and J. Noel. 1991. Do inventory disclosures predict sales and earnings? *Journal of Accounting, Auditing & Finance* 6: 145–181.

Botosan, C. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72 (3): 323–349.

Bryan, S. 1997. Incremental information content of required disclosures contained in management discussion and analysis. *The Accounting Review* 72 (2): 285–302.

Cole, C., and C. Jones. 2004. The usefulness of MD&A disclosures in the retail industry. *Journal of Accounting, Auditing & Finance* 19 (4): 361–388.

Core, J. 2001. A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31 (1–3): 441–456 Discussion.

———, R. Holthausen, and D. Larcker. 1999. Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics* 51 (3): 371–406.

Davis, A., J. Piger, and L. Sedor. 2008. Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases. Working paper, University of Oregon and The University of Washington.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *The Journal of Finance* 46 (4): 1325–1355.

Dye, R. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23 (1): 123–145.

———. 1986. Proprietary and nonproprietary disclosures. *The Journal of Business* 59 (2): 331–366.

———. 1988. Earnings management in an overlapping generations model. *Journal of Accounting Research* 26 (2): 195–235.

———. 2001. An evaluation of "essays on disclosure" and the disclosure literature in accounting. *Journal of Accounting and Economics* 31 (1–3): 181–235.

Fields, T., T. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31 (1–3): 255–307.

Frankel, R., W. Mayew, and Y. Sun. 2009. Do pennies matter? Investor relations consequences of small negative earnings surprises. *Review of Accounting Studies*.

Hanlon, M., S. Rajgopal, and T. Shevlin. 2003. Are executive stock options associated with future earnings? *Journal of Accounting and Economics* 36 (1–3): 3–43.

Healy, P., and K. Palepu. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31 (1–3): 405–440.

———, and J. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (4): 365–384.

Hirst, E., and P. Hopkins. 1998. Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research* 36 (Supplement): 47–75.

Hribar, P., and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (1): 105–134.

Hunton, J., R. Libby, and C. Mazza. 2006. Financial reporting transparency and earnings management. *The Accounting Review* 81 (1): 135–157.

Jo, H., and Y. Kim. 2007. Disclosure frequency and earnings management. *Journal of Financial Economics* 84 (2): 561–590.

Lang, M., and R. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71 (4): 467–492.

Lev, B., and R. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* 31 (2): 190–215.

Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45 (2–3): 221–247.

Lobo, G., and J. Zhang. 2001. Disclosure quality and earnings management. *Asia-Pacific Journal of Accounting and Economics* 8 (1): 1–20.

Mulford, C., and E. Comiskey. 1996. *Financial Warnings*. New York, NY: Wiley.

———, and ———. 2002. *The Financial Numbers Game: Detecting Creative Accounting Practices*. New York, NY: Wiley.

Penman, S. 2001. *Financial Statement Analysis and Security Valuation*. New York, NY: McGraw-Hill Companies.

Revsine, L., D. Collins, and W. B. Johnson. 2004. *Financial Reporting and Analysis*. 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Richardson, V.. 2000. Information asymmetry and earnings management: Some evidence. *Review of Quantitative Finance and Accounting* 15 (4): 325–347.

Rogers, K., and J. Grant. 1997. Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis* 3 (1): 17–30.

Roychowdhury, S. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42 (3): 335–370.

Securities and Exchange Commission (SEC). 1987. *Concept Release on Management's Discussion and Analysis of Financial Condition and Results of Operations*. Securities Act Release No. 6711. Washington, D.C.: SEC.

———. 2003. *Interpretation: Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations*. Securities Act Release No. 8350. Washington, D.C.: SEC.

———. 2005. *SEC Charges Kmart's Former CEO and CFO with Financial Fraud*. Accounting and Auditing Enforcement Release No. 2295. Washington, D.C.: SEC.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–316.

Thomas, J. K., and H. Zhang. 2002. Inventory changes and future returns. *Review of Accounting Studies* 7 (2–3): 163–187.

Trueman, B., and S. Titman. 1988. An explanation for accounting income smoothing. *Journal of Accounting Research* 26 (3): 127–139.

Verrecchia, R. E. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5: 179–194.

———. 2001. Essays on disclosure. *Journal of Accounting and Economics* 31 (1–3): 97–180.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1441–1471

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1441

Cost Behavior and Analysts' Earnings Forecasts

Dan Weiss
Tel Aviv University

ABSTRACT: This study examines how firms' asymmetric cost behavior influences analysts' earnings forecasts, primarily the accuracy of analysts' consensus earnings forecasts. Results indicate that firms with stickier cost behavior have less accurate analysts' earnings forecasts than firms with less sticky cost behavior. Furthermore, findings show that cost stickiness influences analysts' coverage priorities and investors appear to consider sticky cost behavior in forming their beliefs about the value of firms. This study integrates a typical management accounting research topic, cost behavior, with three standard financial accounting topics (namely, accuracy of analysts' earnings forecasts, analysts' coverage, and market response to earnings surprises).

Keywords: *analysts; earnings forecasts; sticky costs; asymmetric cost behavior.*

JEL Classifications: *M41; G12.*

I. INTRODUCTION

Management accountants have traditionally focused on cost behavior as an important aspect of profit analysis for managers. Financial analysts, however, estimate firms' future costs in the process of forecasting future earnings. Predicting cost behavior is, therefore, an essential part of earnings prediction. Yet, a potential relationship between firms' cost behavior and properties of analysts' earnings forecasts has not yet been explored. This study integrates the management and financial accounting disciplines by showing effects of cost behavior on: (1) the accuracy of analysts' consensus earnings forecasts, (2) the extent of analyst coverage, and (3) the market response to earnings announcements.

Focusing on cost behavior, I build on the concept of sticky costs (Anderson et al. 2003). Costs are termed sticky if they increase *more* when activity rises than they decrease when activity falls by an equivalent amount. A firm with stickier costs shows a greater decline in earnings when the

The author is grateful for valuable comments and constructive suggestions from Matthias Amen, Yakov Amihud, Eli Amir, Ramji Balakrishnan, Ray Ball, Rajiv Banker, Avraham Beja, Simon Benninga, Jeffrey Callen, Clara Chen, Harry Evans (editor), Richard Frankel, Dan Givoly, Thomas Hemmer, Surya Janakiraman, Avner Kalay, Itay Kama, Michael Kinney, Thomas Lys, Michael Maher, Aharon Ofer, Naomi Soderstrom, Paul Zarowin, and two anonymous referees. Comments from the participants of the MAS Midyear Conference, Global Management Accounting Research Symposium (GMARS), Manufacturing Accounting Research Conference, the Tel Aviv Accounting Conference, and American Accounting Association Annual Meetings are highly appreciated. Part of this study was performed during the author's visit to the University of California, Davis. The research was supported by the Israel Science Foundation.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: September 2008
Accepted: January 2010
Published Online: June 2010

activity level falls than a firm with less sticky costs. The reason is that stickier costs result in a smaller cost adjustment when activity level declines and, therefore, lower cost savings. Lower cost savings result in a greater decrease in earnings. This greater decrease in earnings when the activity levels fall increases the variability of the earnings distribution, resulting in less accurate earnings predictions.

Results, based on a sample of 44,931 industrial firm quarters for 2,520 firms from 1986 through 2005, indicate that sticky cost behavior reduces the accuracy of analysts' consensus earnings forecasts after controlling for environmental uncertainty, the amount of available firm-specific information, the forecast horizon, and industry effects.

Classifying costs into sticky and anti-sticky costs,[1] findings show that analysts' absolute consensus earnings forecasts for firms with sticky cost behavior are, on average, 25 percent less accurate than those for firms with anti-sticky cost behavior. Evidently, cost behavior is an influential determinant of analysts' forecast accuracy. The results are robust to potential managerial discretion that might bias the cost stickiness measure and to estimating cost stickiness over a long time window. The findings extend Banker and Chen (2006), who show that cost behavior explains a considerable part of analysts' advantage over time-series models. These findings are also useful for investors who use consensus earnings forecasts to value firms, as they suggest that stickier costs indicate more volatile future earnings.

Addressing the extent of analyst coverage, I examine the relationship between the accuracy of earnings forecasts and the extent of analyst coverage. While Alford and Berger (1999) and Weiss et al. (2008) document a positive relationship, Barth et al. (2001) report that analysts tend to prefer covering firms with intangible assets characterized by volatile performance. Thus, prior evidence is mixed and this relationship is an open empirical issue. I find that firms with stickier costs (and less accurate earnings forecasts) have lower analyst coverage after controlling for the amount of available information, environmental uncertainty, intensity of R&D expenditures, and additional determinants of supply and demand for analysts' forecasts reported in the literature (e.g., Bhushan 1989; Lang and Lundholm 1996). Findings indicate that firms' cost behavior affects analysts' coverage priorities.

Finally, I examine whether investors behave as if they understand cost stickiness in responding to earnings announcements. As earnings predictability decreases, reported earnings provide less useful information for the prediction of future earnings, such that the earnings response coefficient decreases (e.g., Lipe 1990). If investors recognize cost stickiness to some extent, being aware that cost stickiness diminishes the accuracy of the analysts' earnings forecasts, then stickier cost behavior causes investors to rely less on realized earnings information because of its lower predictive power. Similarly, I find a weaker market response to earnings surprises for firms with stickier cost behavior. Overall, findings indicate that cost behavior matters in forming investors' beliefs regarding the value of the firm.

This empirical examination is facilitated by a new measure of cost stickiness at the firm level developed in this study. I estimate the difference in cost function slopes between upward and downward activity adjustments. While Anderson et al. (2003) and subsequent studies use cross-sectional and time-series regressions to estimate cost stickiness,[2] the new measure developed here puts less demand on the data and allows for testing the sensitivity of results to key cost model assumptions. The new measure corroborates prior evidence on variation among firms' cost sticki-

[1] Costs are termed anti-sticky if they increase *less* when activity rises than they decrease when activity falls by an equivalent amount. See examples in Balakrishnan et al. (2004) and the discussion in Section II.

[2] See, for instance, Banker et al. (2008) and Anderson and Lanen (2007).

ness and provides room for estimating cost stickiness of firms operating in industries with a small number of firms, which limits a meaningful estimation of regression models.[3]

This study expands the scope of cost behavior research. Traditionally, cost behavior has attracted the attention of management accountants interested in decision-making and control. The current results show that financial analysts benefit from understanding cost behavior as well. Further, these findings contribute to our understanding of how analysts use public information reported in financial statements to recognize cost behavior (e.g., Abarbanell and Bushee 1997; Brown et al. 1987).

In sum, this study integrates a typical management accounting research topic, cost behavior, with three standard financial accounting topics. The importance of integrating the two streams of research has long been recognized (Hemmer and Labro 2008).

Hypotheses are developed in Section II, the research design is described in Section III, and the empirical results are in Section IV. Section V offers a concluding remark on the prospects of integrating management and financial accounting research.

II. DEVELOPMENT OF HYPOTHESES

Despite the wide interest in analysts' earnings forecasts, prior research has not yet investigated the relationship between firms' cost behavior and properties of analysts' earnings forecasts, notwithstanding the essential part that cost prediction plays in the process of earnings prediction. Prior empirical studies provide evidence that the accuracy of analysts' earnings forecasts increases in the amount of information available regarding the firm (Atiase 1985; Lang and Lundholm 1996), increases in firm size but not in firm complexity (Brown et al. 1987), and decreases in the level of uncertainty in the firm's production environment (Parkash et al. 1995). Later work by Banker and Chen (2006) reports that cost behavior explains a considerable portion of the analysts' advantage in earnings prediction over various time-series models.

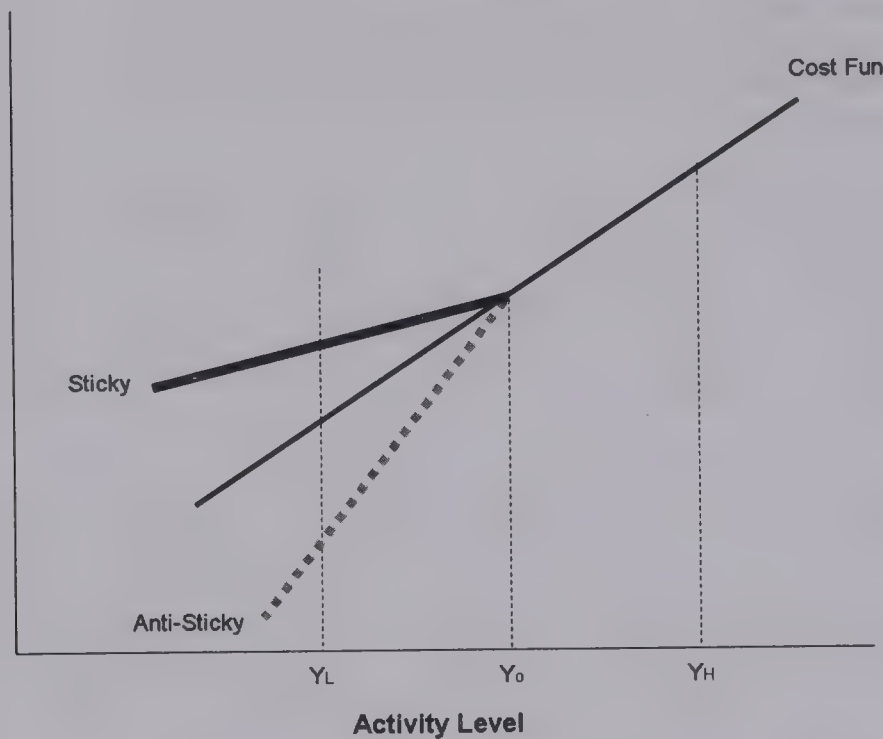
The recently developed concept of sticky costs provides a compelling setting for exploring why and how cost behavior affects the accuracy of analysts' earnings forecasts. Sticky costs indicate that costs tend to "stick" and hence do not go away when activity levels decline. Balakrishnan and Gruca (2008) report that hospital administrators find it hard and expensive to adjust capacity level of core activities downward resulting in sticky costs. Banker and Chen (2006) improve earnings predictions by estimating the excessive costs incurred due to costs being sticky when sales decrease. Anderson et al. (2007) report that cost stickiness causes the ratio of SG&A costs to sales to increase, rather than decrease proportionally with sales, when revenue declines. Overall, prior studies perceive costs as sticky if firms incur disproportionate costs when activity levels decline.

I build on Balakrishnan et al. (2004) to illustrate the intuition underlying the relationship between the extent of cost stickiness and the accuracy of analysts' earnings forecasts. Balakrishnan et al. (2004) argue that the level of capacity utilization affects managers' response to a change in activity level. Thus, if capacity utilization is high, the firm's managers are not likely to immediately cut resources in response to a decrease in activity level because the decrease may be temporary. However, an increase in activity level under high-capacity utilization is likely to cross the available resource threshold and trigger an increase in resources supplied. Assuming high-capacity utilization, the response to a decrease in activity level will be smaller than the response to a similar increase in activity level, resulting in sticky costs—depicted by the bold line in Figure 1.

By contrast, suppose the same firm experiences excess capacity. Its managers are likely to use the slack to absorb the demand from an increase in activity level. However, an additional decrease

[3] For instance, Banker and Chen (2006) exclude from their sample four-digit SIC code industries with less than 20 firms.

FIGURE 1
Cost Asymmetry



The figure depicts sticky and anti-sticky cost functions based on Balakrishnan et al.'s (2004) example. The bold cost function illustrates sticky costs assuming that activity level Y_0 is high-capacity utilization. The dashed cost function illustrates anti-sticky costs assuming excess capacity for activity level Y_0.

in activity level is interpreted as confirming a permanent reduction in demand and triggers a response. Assuming excess capacity, the cost response to an activity level decrease exceeds the cost response to a similar increase in activity level, resulting in anti-sticky costs—depicted by the dashed line in Figure 1.

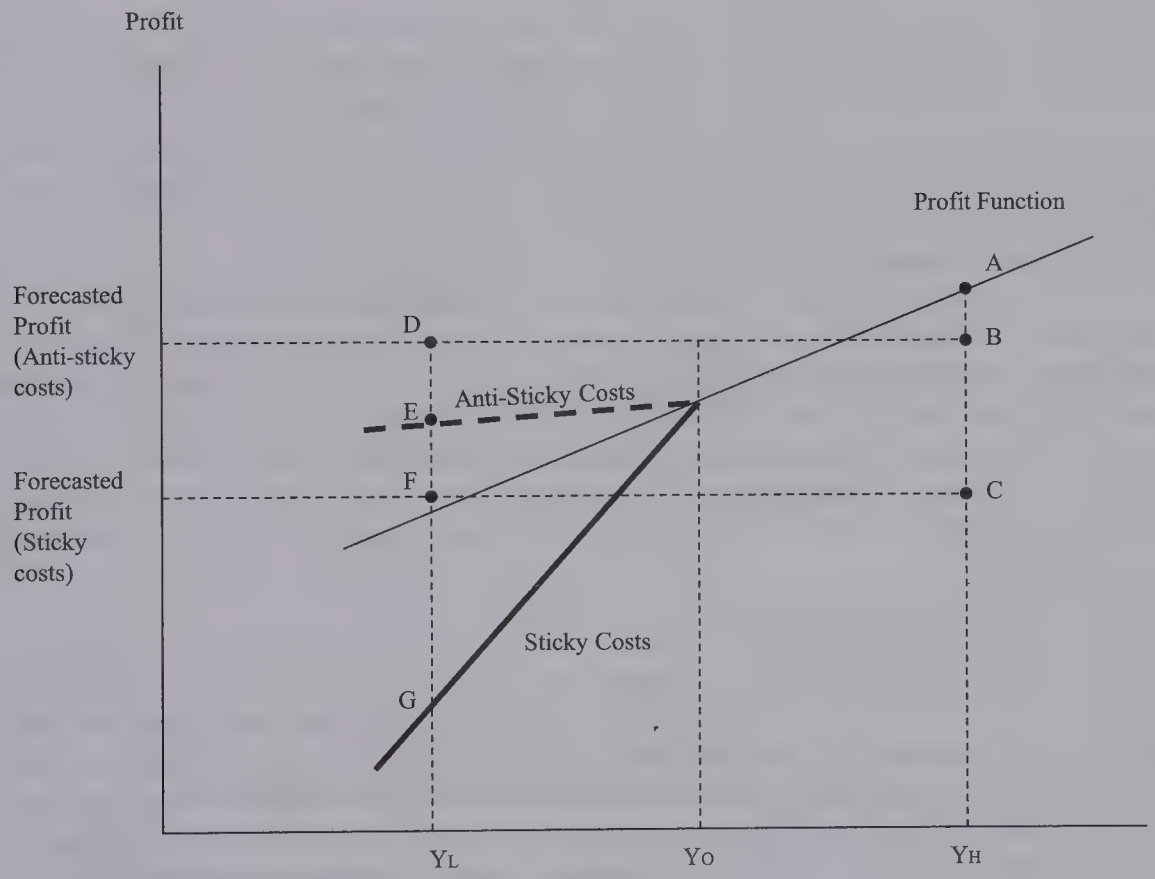
Next, I build on Balakrishnan et al. (2004) to illustrate that stickier costs result in greater earnings variability. Higher capacity utilization yields stickier costs, resulting in lower cost savings and a greater decrease in profits when sales decrease.[4] Assuming that the rest of the distribution of profits is unchanged, this greater decrease in profits increases the variability of the *ex ante* profit distribution.

Now, suppose an analyst predicts future profits. For simplicity, I assume that future activity level will either increase or decrease by an equivalent amount with equal probability. I further suppose that the analyst recognizes cost behavior to a reasonable extent and assume that the analyst forecasts expected profit (e.g., Ottaviani and Sorensen 2006). Other things equal, sticky costs result in lower profits when the activity level declines than anti-sticky costs do. Assuming equal profits when activity levels rise, the analyst's profit forecast is lower under sticky costs than

[4] The terms profits and earnings are used interchangeably in this study.

under anti-sticky costs. For that reason, the *absolute* forecast error when activity levels decline as well as when activity levels rise is greater under sticky costs than under anti-sticky costs. Figure 2 depicts lower profits under sticky costs (assuming high-capacity utilization) than under anti-sticky costs (assuming excess capacity) when activity levels decline: point G is below point E. Lower profits are expected and forecasted under sticky costs than under anti-sticky costs; see profit levels FC and DB, respectively. The *absolute* forecast errors under sticky costs are larger than under anti-sticky costs: AC > AB when the activity level increases and FG > DE when the activity level decreases. The example demonstrates that the absolute forecast errors increase with the extent of cost stickiness.

FIGURE 2
Absolute Forecast Errors are Greater in the Presence of Sticky Costs Than in the Presence of Anti-Sticky Costs



The figure depicts two profit functions following Balakrishnan et al. (2004). The bold profit function assumes sticky costs and the dashed profit function assumes anti-sticky costs. When the activity level declines from Y_0 to Y_L, profits are lower under sticky costs (assuming high-capacity utilization) than under anti-sticky costs (assuming excess capacity): point G is below point E. Expected profits are lower under sticky costs than under anti-sticky costs: see profit levels FC and DB, respectively. The *absolute* forecast errors under sticky costs are larger than under anti-sticky costs: AC > AB on activity level increases and FG > DE on activity level decreases. Overall, the *absolute* forecast errors when activity levels decline, as well as when activity levels increase, are greater under sticky costs than under anti-sticky costs.

The above example is used to illustrate the intuition of the relationship between the extent of cost stickiness and the accuracy of earnings forecasts. This relationship between the extent of cost stickiness and the absolute forecast errors is modeled in the Appendix (consistent with Equation (5) in Banker and Chen [2006]). The Proposition presented in the Appendix indicates a positive relationship between the extent of cost stickiness and the absolute forecast errors. Accordingly, my first hypothesis is:[5]

H1: Increased cost stickiness reduces the accuracy of analysts' consensus earnings forecasts.

Prior literature documents a relationship between the accuracy of analysts' earnings forecasts and the extent of analyst coverage (e.g., Alford and Berger 1999). Recently, Weiss et al. (2008, Table 7) report that firms with high analyst coverage have more accurate earnings forecasts than firms with low analyst coverage. Stickel (1992) reports that members of the Investor All-American Research Team have more accurate forecasts than non-members. Analysts who find this competition to be of major importance are likely to prefer covering firms with less sticky cost behavior to achieve greater expected accuracy.[6]

However, Barth et al. (2001) report high coverage of firms with intangible assets, characterized by low earnings predictability and high earnings forecasts errors. While analysts are motivated to provide investors with more accurate earnings forecasts, they may not shy away from following a firm with low earnings predictability if they have an information advantage with respect to that firm or if the demand for forecasts is higher for that firm. In sum, the empirical evidence on the relationship between the accuracy of analysts' earnings forecasts and the extent of analyst coverage is mixed, and it remains an open empirical issue.

I examine whether sticky cost behavior influences the extent of analyst coverage. Sticky cost behavior will influence analysts' coverage priorities if they recognize the relationship between cost stickiness and accuracy of earnings forecasts hypothesized above. I test a potential relationship between sticky cost behavior and the extent of analyst coverage after controlling for the intensity of research and development, the amount of available information, firm size, environmental uncertainty, and for additional determinants of supply and demand for analysts' forecasts reported in the literature. Because there is no *ex ante* basis for a prediction, the corresponding hypothesis is stated in the null form:

H2: Sticky cost behavior does not affect analyst coverage.

Finally, I examine whether investors recognize cost behavior. If investors have some understanding that firms with stickier costs tend to have less accurate earnings forecasts, then cost behavior is likely to influence their response to surprises in earnings announcements. As earnings predictability decreases, reported earnings provide less useful information for valuation and prediction of future earnings, resulting in a lower earnings response coefficient (e.g., Lipe 1990). Abarbanell et al. (1995) show that the earnings-price response coefficient increases in the forecast precision. If investors recognize that cost stickiness diminishes the accuracy of the analyst's

---

[5] To clarify, several studies report a variation in the level of cost stickiness—see Balakrishnan and Gruca (2008), Banker et al. (2008), Anderson and Lanen (2007), and Balakrishnan and Soderstrom (2009). These studies offer explanations for both sticky and anti-sticky cost behavior (e.g., ownership, pessimism/optimism with respect to future demand, core activity). In this study, I build on this variation and show a relationship between the level of cost stickiness and properties of analysts' earnings forecasts.

[6] Examining factors associated with the extent of analyst coverage, Bhushan (1989) finds that the number of analysts covering a firm increases in firm size, O'Brien and Bhushan (1990) report greater coverage in industries with stringent disclosure requirements, and Lang and Lundholm (1996) claim greater coverage of firms with more informative disclosure policies. Frankel et al. (2006), however, find no relation between the informativeness of the analysts' forecasts and the size of the analyst following.

earnings forecasts, then stickier cost behavior causes investors to rely less on realized earnings information because of its low predictive power. The third hypothesis summarizes the argument:

H3: The market response to earnings surprises is weaker for firms with stickier cost behavior.

If H3 holds, then it suggests that investors have some appreciation of the role of cost behavior in determining earnings surprises. In other words, this hypothesis predicts that cost behavior matters in forming investors' beliefs regarding the value of firms.

III. RESEARCH DESIGN

Focusing on asymmetric cost behavior, this study proposes a new measure of cost stickiness at the firm level. Prior studies use a cross-sectional regression model to estimate cost stickiness at the industry level or a time-series regression model to estimate it at the firm level (e.g., Anderson et al. 2003). Taking a different path, this study introduces a direct measure of cost stickiness at the firm level. I estimate the difference between the rate of cost decrease for recent quarters with decreasing sales and the corresponding rate of cost increase for recent quarters with increasing sales:

$$STICKY_{i,t} = \log\left(\frac{\Delta COST}{\Delta SALE}\right)_{i,\underline{\tau}} - \log\left(\frac{\Delta COST}{\Delta SALE}\right)_{i,\bar{\tau}} \quad \underline{\tau},\bar{\tau} \in \{t, \ldots, t-3\},$$

where $\underline{\tau}$ is the most recent of the last four quarters with a decrease in sales and $\bar{\tau}$ is the most recent of the last four quarters with an increase in sales, $\Delta SALE_{it} = SALE_{it} - SALE_{i,t-1}$, (Compustat #2), $\Delta COST_{it} = (SALE_{it} - EARNINGS_{it}) - (SALE_{i,t-1} - EARNINGS_{i,t-1})$, and *EARNINGS* is income before extraordinary items (Compustat #8).

STICKY is defined as the difference in the cost function slope between the two most recent quarters from quarter $t-3$ through quarter t, such that sales decrease in one quarter and increase in the other. If costs are sticky, meaning that they increase *more* when activity rises than they decrease when activity falls by an equivalent amount, then the proposed measure has a negative value. A *lower* value of *STICKY* expresses *more* sticky cost behavior.[7] That is, a negative (positive) value of *STICKY* indicates that managers are less (more) inclined to respond to sales drops by reducing costs than they are to increase costs when sales rise.

Following prior sticky costs studies, *STICKY* uses a change in sales as an imperfect proxy for the actual activity change because changes in activity level are not observable. Employing sales as a fundamental stochastic variable is in line with Dechow et al. (1998), who suggest a model of earnings, cash flow, and accruals, assuming a random walk sales process. Banker and Chen (2006) also use sales as a fundamental stochastic variable for predicting future earnings.

Since analysts estimate total costs in the process of earnings prediction, the stickiness measure concentrates on total costs to gain insights into a potential relationship between stickiness of total costs and the accuracy of analysts' earnings predictions. Investigating how cost stickiness affects analysts' earnings forecasts, I use sales minus earnings. Employing total costs for the analysis also eliminates managerial discretion in cost classifications (Anderson and Lanen 2007). I also assume that costs increase in activity level (as in the adjustment costs model presented in the Appendix). This assumption means that costs move in the same direction as activity and precludes cost increases when activity falls and cost decreases when activity increases (Anderson and Lanen 2007). For this reason, I do not use observations with costs that move in opposite directions in estimating *STICKY*. The ratio form and logarithmic specification make it easier to compare vari-

[7] The estimate of *STICKY* is consistent with the sign of the parameter α, as defined in the model presented in the Appendix. The sign of *STICKY* is also consistent with the stickiness measure, β_2, in Anderson et al. (2003, Model I).

ables across firms, as well as alleviating potential heteroscedasticity (Anderson et al. 2003).

The proposed measure has several advantages. First, and most important for this study, *STICKY* estimates cost asymmetry at the firm level. Thus, it provides means of investigating how cost behavior influences analysts' earnings forecasts. Moreover, it allows for a large-scale study without restricting the analysis to firms with at least ten valid observations and at least three sales reductions during the sample period (see Anderson et al. 2003, 56).[8]

Second, by design, the stickiness of a linear cost function is zero, i.e., $STICKY = 0$, for a traditional fixed-variable cost model with a constant slope for all activity levels within a relevant range. Thus, a zero value indicates that managers change costs symmetrically in response to sales increases and declines.

Third, the proposed cost stickiness measure has a wider scope than that used by Anderson et al. (2003) because it accounts for a difference between proportions of cost changes and allows for cost friction with respect to sales increases. For instance, Chen et al. (2008, 2) argue that empire-building incentives are "likely to lead managers to increase SG&A costs too rapidly when demand increases." They report a positive association between managerial empire building incentives and the degree of cost asymmetry. *STICKY* affords an examination of how cost asymmetry affects the forecast accuracy in the presence of decreases in sales (i.e., as presented by Anderson et al. [2003]) and in the presence of increases in sales.

Nonetheless, there are potential measurement errors in the suggested cost stickiness metric. First, the model assumes a piecewise linear specification of the cost function within the relevant range of activity, which simplifies the analysis and allows for measuring cost stickiness when the upward and downward activity changes do not have the same magnitude. This approximation is consistent with prior studies on sticky costs and is reasonable in the context of investigating a relationship between attributes of cost behavior and properties of analysts' forecasts.

Second, the model assumes a realization of an exogenous state of the world that determines activity level. However, growth or reduction in activity can occur not only because of changes in activity level, but also because of changes in prices of products or resources or other managerial choices (Anderson and Lanen 2007). I restrict the sample to competitive industrial firms to minimize this problem, and later test the sensitivity of results to potential managerial discretion.

To check consistency with prior literature, I compute the suggested measure for two major cost categories investigated in prior literature. Specifically, *COGS-STICKY* and *SGA-STICKY* substitute changes in total costs with changes in cost of goods sold, hereafter *COGS* (Compustat #30) and *SGA* costs (Compustat #1), respectively. The median proportion of *COGS* and *SGA* to sales in my sample is 64.7 percent and 23.1 percent, respectively. However, the accounting classification of *COGS* and *SGA* is open to managerial judgment, which may introduce bias into the cost stickiness estimate of specific cost components. The results should be interpreted in light of this limitation. Taken as a whole, the stickiness measure is expected to provide broad insights into the relationship between cost behavior and properties of analysts' earnings forecasts.

Measuring the accuracy of the analyst consensus forecast, I employ the mean absolute earnings forecast errors as an inverse accuracy measure. This accuracy gauge has been extensively used in the accounting literature (e.g., Lang and Lundholm 1996). Thus, the forecast error is defined as:

$$FE_{it} = \frac{actual\ EPS_{it} - analyst\ consensus\ forecast_{it}}{Price_{i,t-1}},$$

and the absolute forecast error is $ABS\text{-}FE_{it} = |FE_{it}|$, where the analyst consensus forecast is the

[8] In measuring skewness of firm-specific earnings distributions, Gu and Wu (2003) require each firm to have at least 16 quarterly observations.

mean of analyst forecasts for firm i and quarter t announced in the month immediately preceding that of the earnings announcement. The relatively narrow time window and the short forecast horizon control for the timeliness of the forecasts and mitigate a potential trade-off between timing and accuracy (Clement and Tse 2003).

Testing H1

In testing whether stickier cost behavior results in greater mean absolute analyst consensus earnings forecast error, I control for the amount of available firm-specific information, the inherent uncertainty in the operations environment, and the forecast horizon. The literature reports that an increased amount of available firm-specific information reduces the forecast error. The amount of information acquired by analysts is positively related to firm size (Atiase 1985; Collins et al. 1987; Bhushan 1989). Accordingly, I use firm size as a control variable and expect a negative coefficient. Brown (2001) reports a disparity between the magnitude of earnings surprises of profits and losses. I use an indicator variable to control for losses because they reflect more timely information and are associated with larger absolute forecast errors than are profits. I also follow Matsumoto (2002) and control for potential earnings guidance, which is likely to reduce the forecast error if it results in meeting or slightly beating the consensus earnings forecast.

Environmental uncertainty is likely to influence the forecast accuracy. If the business environment is highly volatile, then one will expect larger forecast errors. I use two proxies for the level of environmental uncertainty. The first is the coefficient of variation in sales, which directly captures sales volatility. The second is analyst forecast dispersion, which measures complementary uncertainty aspects of firms' earnings (Barron et al. 1998). Brown et al. (1987) and Wiedman (1996) report that the accuracy of analysts' forecasts decreases in the dispersion of analysts' forecasts, which is used to proxy for the variance of information observations.

In addition, management accounting textbooks (e.g., Maher et al. 2006) present cost-volume-profit analysis and suggest that firms with a high operating leverage are likely to exhibit high earnings volatility. Lev and Thiagarajan (1993) employ profit margin as a proxy for operating leverage. In an early study, Adar et al. (1977) present a positive relationship between profit margin and forecast error in a cost-volume-profit under uncertainty setting. Operating leverage varies across firms and is likely to depend on the firm-specific business environment, as well as current macroeconomic conditions. The higher the operating leverage of the firm, the higher is the expected error in the analysts' earnings forecast. Therefore, I predict a positive relationship between operating leverage and the magnitude of analysts' earnings forecast errors.

I also control for unexpected contemporaneous seasonal shocks to earnings. An indicator variable, *SEASON*, indicates firm quarters with a positive change in earnings from the same quarter in the prior year. This variable controls for the relation between the change in earnings and the forecast error (Matsumoto 2002). A positive coefficient estimate is predicted.

I estimate the following three cross-sectional regression models with two-digit SIC code industry effects:

Model 1(a)

$$ABS\text{-}FE_{it} = \beta_0 + \beta_1 STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it};$$

Model 1(b)

$$ABS\text{-}FE_{it} = \beta_0 + \beta_1 COGS\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it};$$

Model 1(c)

$$ABS\text{-}FE_{it} = \beta_0 + \beta_1 SGA\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it};$$

where:

MV_{it} = log of market value of equity (Compustat #61 × #14) at quarter end;
$LOSS_{it}$ = indicator variable that equals 1 if the reported earnings (Compustat #8) are negative, and 0 otherwise;
$DOWN_{it}$ = as defined in Matsumoto (2002) and equals 1 if unexpected earnings forecasts are negative, and 0 otherwise;
$VSALE_{it}$ = coefficient of variation of sales measured over four quarters from $t-3$ through t;
$DISP_{it}$ = standard deviation of the analysts' forecasts announced for firm i and quarter t in the month immediately preceding that of the earnings announcement, deflated by stock price at the end of quarter $t-1$;
$OPLEV_{it}$ = ratio between $SALE_{it}$, minus *COGS* (Compustat #30) and $SALE_{it}$; values below 0 or above 1 are winsorized; and
$SEASON_{it}$ = indicator variable that equals 1 if the change in earnings from the same quarter in the prior year (Compustat #8) is positive, and 0 otherwise.

If the above metric captures cost stickiness, then H1 predicts $\beta_1 < 0$ in all three models, where lower values of *STICKY* (*COGS-STICKY*, *SGA-STICKY*) indicate stickier cost behavior.

I further test the sensitivity of the results to the model's assumptions and potential measurement errors in three ways. First, I test the sensitivity of the cost stickiness measure to a longer time window. I compute the ratio of change in total costs to change in sales using data from the most recent eight quarters, $t-7$ through t. I then estimate *M-STICKY*, that is, the difference between the mean slope under downward adjustments and the mean slope under upward adjustments. Thus, *M-STICKY* accounts for downward adjustments and upward adjustments made over eight quarters. Comparing *M-STICKY* with *STICKY* provides insights into the perseverance of firms' cost behavior over a longer window. To check the robustness of the coefficient estimates, I estimate the following regression model:

Model 1(d)

$$ABS\text{-}FE_{it} = \beta_0 + \beta_1 M\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it}.$$

Again, if the above metric captures the cost stickiness, then H1 predicts $\beta_1 < 0$ in model 1(d).

Second, I conduct a limited examination of the effects of cost stickiness generated by past decisions, such as technology choice and labor compensation contracts, on absolute forecast errors. Specifically, I consider two forms of managerial discretion: current decisions made in response to realized market conditions in the current quarter t, and past decisions made over quarters prior to quarter t. I view adjustments of activity levels as responses in reaction to realized market conditions, in contrast to prior decisions. Substituting $STICKY_{i,t-1}$ for $STICKY_{i,t}$ allows for estimating the impact of past decisions only. In other words, $STICKY_{i,t-1}$ proxies for the extent of cost stickiness in an earlier quarter, excluding all managerial discretionary choices made in quarter t, such as price discounts or accrual manipulations.

Model 1(e)

$$ABS\text{-}FE_{it} = \beta_0 + \beta_1 STICKY_{i,t-1} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it}.$$

As before, the hypothesis predicts $\beta_1 < 0$ in model 1(e).

Third, I collect evidence concerning the assumption that analysts understand cost behavior to some extent. If analysts ignore cost stickiness when it exists, then their earnings forecasts will be upward biased. Similarly, if analysts ignore anti-sticky costs when they exist, then their earnings forecasts will be downward biased. However, if analysts understand cost behavior, then the mean forecast error (not absolute error) for firms with sticky costs as well as for firms with anti-sticky costs should *not* be affected by the level of cost stickiness.[9] In other words, if analysts recognize cost behavior, then the extent of cost stickiness will not influence the mean signed forecast error.

Testing H2

To test the association between cost stickiness and analyst coverage, I regress the number of analysts following a firm on its cost stickiness and control variables. The analyses include independent variables to control for the amount of available information, environmental uncertainty, the intensity of research and development expenditures, additional determinants of supply and demand for analysts' forecasts reported in the literature, and year effects and two-digit SIC code industry effects.

Prior literature reports that firm size is a primary determinant of analyst coverage (Bhushan 1989; Hong et al. 2000; Das et al. 2006), perhaps because large firms have more available firm-specific information than small firms (Collins et al. 1987). The extent of information asymmetry between managers and investors is likely to enhance the demand for earnings forecasts, but analysts are required to invest more resources in acquiring information. I use research and development expenditures as a proxy for information asymmetry because firms with more intangible assets exhibit greater information asymmetry (Barth et al. 2001; Barron et al. 2002).

Controlling for uncertainty in the forecasting environment, I employ the coefficient of variation in sales as a direct measure for shocks in demand. In addition, analyst forecast dispersion and the absolute forecast error in the prior quarter are included to measure other aspects of the uncertainty in firms' earnings (Brown et al. [1987] and Matsumoto [2002], respectively). Das et al. (1998) argue that analysts extract higher rents by following less predictable firms, because demand for private information is the highest for these firms, but the accuracy of the forecasts is expected to be lower. Thus, the net effect of uncertainty in the forecasting environment on an increase in the extent of analyst following is ambiguous.

I also control for growth and trading volume (Lang and Lundholm 1996), which provide analysts with greater incentives to cover firms. Finally, Baik (2006) argues that firms experiencing financial distress appear to suffer from self-selection by analysts. Accordingly, I also control for losses.

Using count-data in the dependent variable, I follow Rock et al. (2000) and use the standard negative binomial distribution to estimate regression models 2(a) through 2(c):

[9] This argument recognizes that analysts announce expected earnings as their forecast. Even if analysts' forecasts are biased (say, optimistically), there is no reason to believe that their bias depends on the level of cost stickiness.

Model 2(a)

$$FLLW_{it} = \beta_0 + \beta_1 STICKY_{it} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it};$$

Model 2(b)

$$FLLW_{it} = \beta_0 + \beta_1 M\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it};$$

Model 2(c)

$$FLLW_{it} = \beta_0 + \beta_1 STICKY_{i,t-1} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it};$$

where year effects and two-digit SIC code industry effects are added to all models, and:

$FLLW_{it}$ = number of analysts' earnings forecasts announced for firm i and quarter t in the month immediately preceding that of the earnings announcement;
$GROWTH_{it} = (SALE_{it} / SALE_{i,t-4})^{0.25} - 1$;
RD_{it} = Compustat #4 for firm i in quarter t divided by $SALE_{it}$; observations with no values are set equal to 0 and values are winsorized at 1; and
TV_{it} = quarterly trading volume in millions of shares.

Model 2(b) measures cost stickiness based on data from eight preceding quarters and model 2(c) uses a lagged measure of cost stickiness estimated on a prior quarter to strengthen the causality argument.

Market Tests of H3

The third hypothesis predicts that the market response to earnings surprises is weaker for firms with stickier costs than for firms with less sticky costs. To test this hypothesis, I estimate a valuation model that regresses the cumulative abnormal return on the magnitude of earnings surprise and the interaction between earnings surprise and cost stickiness, while controlling for environmental uncertainty. I use contemporaneous estimates of cost stickiness and an earlier quarter estimate in pooled cross-sectional regression models. Additionally:

Model 3(a)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it};$$

Model 3(b)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} COGS\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it};$$

Model 3(c)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} SGA\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it};$$

Model 3(d)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} M\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it};$$

Model 3(e)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} D\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it};$$

Model 3(f)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} STICKY_{i,t-1} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it};$$

where $CAR_{i,t}$ (cumulative abnormal return) is the three-trading-day cumulative value-weighted market-adjusted abnormal return surrounding the earnings announcement for firm i in quarter t, and, $D\text{-}STICKY_{i,t}$ equals -1 if $STICKY_{i,t} < 0$, and 0 otherwise.

To control for environmental uncertainty, Imhoff and Lobo (1992) use dispersion in analysts' earnings forecasts and show that firms with higher *ex ante* earnings uncertainty exhibit smaller price changes in response to earnings announcements than firms with lower *ex ante* earnings uncertainty. Dispersion in analysts' earnings forecasts is likely to capture additional aspects of earnings predictability other than those related to cost behavior. To control for broad aspects of earnings predictability, I employ both the dispersion in analysts' forecasts, *DISP*, and *VSALE* as control variables in the above models. If the coefficient estimate for the interaction variable becomes statistically significant after controlling by *DISP* and *VSALE*, then this suggests that cost behavior matters in forming investors' beliefs regarding the value of the firm. The earnings response coefficient is predicted to be weaker for firms with stickier cost behavior (i.e., $\beta_2 > 0$).

Sample Selection

The sample includes all industrial firms (SIC codes 2000–3999) from 1986 to 2005. The study is limited to industrial firms for two reasons. First, it allows examination of the effects of a potential variation in cost stickiness of the *COGS* and *SGA* cost components on the accuracy of the earnings forecasts. The homogenous structure of the profit and loss statement among industrial firms allows insights into the effects of sticky cost behavior among major cost components on the accuracy of analysts' earnings forecasts. Second, industrial firms (in contrast to utilities and other regulated industries) generally operate in competitive markets, which partially mitigates the measurement error due to a potential pricing effect, rather than to a volume effect.

The data are obtained from Compustat, I/B/E/S, and CRSP. For each firm quarter, I use the consensus forecast calculated as the average of all forecasts announced in the month preceding that of the earnings announcement. Actual earnings are taken from I/B/E/S, as they are more likely to be consistent with the forecast in treating extraordinary items and some special items (Philbrick and Ricks 1991). Following Gu and Wu (2003), I require stock prices to be at least \$3 to avoid the small deflator problem. Announcement dates are taken from Compustat rather than I/B/E/S, which has more firm quarters with missing announcement dates. In line with the model assumption, I limit the sample to firm-year observations, in which costs and sales change in the same direction. This reduces the sample size by 14.1 percent, resulting in a final sample that consists of 44,931 firm quarters for 2,520 firms.

Descriptive Statistics and Consistency with an Earlier Cost Stickiness Measure

Table 1 presents summary statistics for the relevant variables. The mean (median) value of *STICKY* is -0.0174 (-0.0111). Consistent with prior literature, the mean (median) value of *SGA-STICKY* is -0.0306 (-0.0326). Both means are negative and significant ($p < 0.01$). On average, total costs and *SGA* costs exhibit cost stickiness. The mean (median) value of *COGS* is positive, 0.0187 (0.0063). Thus, on average, *SGA* costs exhibit sticky cost behavior, while *COGS*

TABLE 1

Descriptive Statistics, Pooled over Time 1986–2005

Variables	n	Mean	Std. Dev.	Q1	Median	Q3	% Negative
STICKY	44,931	−0.0174	0.4897	−0.1551	−0.0111	0.1205	53.2
COGS-STICKY	37,521	0.0187	0.4707	−0.1564	0.0063	0.1823	48.7
SGA-STICKY	23,809	−0.0306	0.6944	−0.3870	−0.0326	0.3304	55.1
M-STICKY	44,931	−0.0117	0.2398	−0.0633	−0.0094	0.0501	54.9
FE	44,931	−0.0014	0.0382	−0.0110	0	0.0011	38.3
ABS-FE	44,931	0.0071	0.0118	0.0003	0.0011	0.0034	NA
MV	44,926	4.7037	2.1314	3.1508	4.6328	6.1997	NA
LOSS	44,931	0.1628	0.3692	0	0	0	NA
FLLW	44,931	5.5356	5.1759	2	4	8	NA
DOWN	39,415	0.5702	0.4828	0	1	1	NA
VSALE	44,926	0.1480	0.1478	0.0611	0.1026	0.1752	NA
DISP	44,931	0.0018	0.0328	0.0001	0.0004	0.0012	NA
OPLEV	44,559	0.3670	0.1819	0.2408	0.3530	0.4874	NA
SEASON	44,626	0.6111	0.4875	0	1	1	NA
GROWTH	35,864	0.0418	0.1310	0.0142	0.0247	0.0422	NA
RD	44,931	0.0482	0.0987	0	0	0.0620	NA
TV	42,029	5.3360	8.2114	1.8441	2.2254	6.1813	NA

Variable definitions for each firm i on quarter t:

FE_{it} = difference between reported earnings and the mean (consensus) forecasts announced in the month immediately preceding that of the earnings announcement, deflated by the price at the end of the prior quarter. $ABS\text{-}FE_{it} = |FE_{it}|$;

$STICKY_{it} = \log\left(\frac{\Delta COST}{\Delta SALE}\right)_{i,\underline{\tau}} - \log\left(\frac{\Delta COST}{\Delta SALE}\right)_{i,\bar{\tau}}$, $\underline{\tau}$, $\bar{\tau}\varepsilon\{t, \ldots, t-3\}$, where $\underline{\tau}$ is the most recent quarter with sales decrease and $\bar{\tau}$ is the most recent quarter with sales increase;

$COST_{it}$ = sales (Compustat #2) minus net earnings (Compustat #8) for firm i in quarter t;

$SALE_{it}$ = Compustat #2 for firm i in quarter t;

$COGS\text{-}STICKY_{it} = \log\left(\frac{\Delta COGS}{\Delta SALE}\right)_{i,\underline{\tau}} - \log\left(\frac{\Delta COGS}{\Delta SALE}\right)_{i,\bar{\tau}}$, $\underline{\tau}$, $\bar{\tau}\varepsilon\{t, \ldots, t-3\}$, where $\underline{\tau}$ is the most recent quarter with sales decrease and $\bar{\tau}$ is the most recent quarter with sales increase;

$COGS_{it}$ = Compustat #30 for firm i in quarter t;

$SGA\text{-}STICKY_{it} = \log\left(\frac{\Delta SGA}{\Delta SALE}\right)_{i,\underline{\tau}} - \log\left(\frac{\Delta SGA}{\Delta SALE}\right)_{i,\bar{\tau}}$, $\underline{\tau}$, $\bar{\tau}\varepsilon\{t, \ldots, t-3\}$, where $\underline{\tau}$ is the most recent quarter with sales decrease and $\bar{\tau}$ is the most recent quarter with sales increase;

SGA_{it} = Compustat #1 for firm i in quarter t;

$M\text{-}STICKY_{it}$ = difference between the mean cost function slope under upward adjustments made on quarters from $t-7$ through t and the mean cost function slope under downward adjustments made on quarters from $t-7$ through t;

MV_{it} = log of market value of equity (Compustat #61 × #14) on quarter end;

$LOSS_{it}$ = indicator variable that equals 1 if the reported earnings (Compustat #8) are negative, and 0 otherwise;

$FLLW_{it}$ = number of analysts' earnings forecasts announced for firm i and quarter t in the month immediately preceding that of the earnings announcement;

$DOWN_{it}$ = defined in Matsumoto (2002) and equals 1 if unexpected earnings forecasts are negative, and 0 otherwise;

$VSALE_{it}$ = coefficient of variation of sales measured over four quarters from $t-3$ through t;

$DISP_{it}$ = standard deviation of the analysts' forecasts announced for firm i and quarter t during the 30 days

(continued on next page)

TABLE 1 (continued)

prior to the earnings announcement, deflated by the stock price at the end of quarter $t-1$;

$OPLEV_{it}$ = $\frac{\text{SALE}_{it} - \text{COGS}_{it}(\text{Compustat \# 30})}{\text{SALE}_{it}}$, and values below 0 or above 1 are winsorized;

$SEASON_{it}$ = indicator variable that equals 1 if the change in earnings from the same quarter in the prior year (Compustat #8) is positive, and 0 otherwise;

$GROWTH_{it}$ = $(SALE_{it} / SALE_{i,t-4})^{0.25} - 1$;

RD_{it} = Compustat #4 for firm i in quarter t divided by $SALE_{it}$. Observations with no values are taken at 0 and values are winsorized at 1; and

TV_{it} = trading volume in millions of shares.

exhibit anti-sticky cost behavior.[10] The linear nature of raw materials consumption may partially explain this disparity in cost behavior. Another potential explanation for this finding is that salaries and advertising expenses are likely to be classified as *SGA*. The cost stickiness of total costs is also in line with the negative skewness of the earnings distribution reported by Givoly and Hayn (2000) and Gu and Wu (2003). The standard deviation of *STICKY*, *SGA-STICKY*, and *COGS-STICKY* is 0.4897, 0.6944, and 0.4707, respectively, indicating considerable variation among firms' cost behavior.

Examining whether the classification of per firm cost stickiness tends to remain persistent over time, the likelihood of keeping the same cost classification (either sticky or anti-sticky) over two consecutive quarters is 72.5 percent (not tabulated). The Spearman (Pearson) correlation between *STICKY* and *M-STICKY* reported in Table 2 is 0.48 (0.45), indicating reasonable perseverance over eight quarters. Additionally, the Pearson (Spearman) coefficient between the $STICKY_{i,t-1}$ and $STICKY_{it}$ estimates is 0.43 (0.44), both significant at $\alpha = 1$ percent (not tabulated), indicating that firms' cost behavior is reasonably stable over quarters.

As expected, *STICKY* is significantly and positively correlated with both *COGS-STICKY* and *SGA-STICKY*. The correlation between *COGS-STICKY* and *SGA-STICKY* is also positive and significant, indicating a pattern in firms' cost behavior with respect to total costs and to the two cost constituents. The correlation coefficient between *STICKY* and *ABS-FE* is negative and significant, suggesting a negative relation between the cost stickiness and the absolute analysts' earnings forecast errors.

I concentrate on *SGA* costs to check the consistency of the proposed measure with the stickiness measure and results reported by Anderson et al. (2003). I estimate the stickiness of *SGA* costs using the following cross-sectional regression model for two-digit SIC code industries with at least 25 observations:

Model 4

$$\log\left[\frac{SGA_{it}}{SGA_{i,t-1}}\right] = \lambda_0 + \lambda_1 \log\left[\frac{SALE_{it}}{SALE_{i,t-1}}\right] + \lambda_2 SALEDEC_{it} \log\left[\frac{SALE_{it}}{SALE_{i,t-1}}\right] + \varepsilon_{it}$$

where $SALEDEC_{it}$ equals 1 if $SALE_{it} < SALE_{i,t-1}$, and 0 otherwise.

Anderson et al. (2003) suggest the regression coefficient estimate λ_2 as a cost stickiness measure. I compute mean *SGA-STICKY* for two-digit SIC code industries and examine the corre-

[10] Similarly, Anderson and Lanen (2007, Table 7) report that, on average, the number of employees, labor costs, and PPE costs are anti-sticky.

TABLE 2

Correlation Coefficients

Variables	*FE*	*ABS-FE*	*STICKY*	*COGS-STICKY*	*SGA-STICKY*	*M-STICKY*	*MV*	*LOSS*
FE		0.00	0.18**	0.15**	0.09**	0.11**	0.11**	−0.16**
ABS-FE	−0.26**		−0.03**	−0.04**	−0.01	−0.01**	−0.35**	0.28**
STICKY	0.03**	−0.04**		0.48**	0.40**	0.48**	0.02**	−0.17**
COGS-STICKY	0.02**	−0.04**	0.49**		0.07**	0.32**	0.03**	−0.10**
SGA-STICKY	0.01**	−0.03**	0.43**	0.18**		0.12**	0.01	−0.08**
M-STICKY	0.04**	−0.05**	0.45**	0.36**	0.12**		0.00	−0.09**
MV	0.02**	−0.03**	0.01**	0.02**	0.00	0.01		−0.10**
LOSS	−0.04**	0.06**	−0.15**	−0.08**	−0.08**	−0.07**	−0.10**	
FLLW	0.01**	−0.03**	0.02**	0.02**	0.00	0.02	0.74**	−0.09**
DOWN	0.02	−0.03**	0.04**	0.01	0.02	0.04**	0.01	−0.05**
VSALE	−0.01	0.02**	0.00	0.01	0.00	0.00	−0.02**	0.21**
DISP	0.00	0.25**	0.01	0.00	−0.01	0.00	−0.02**	0.16**
OPLEV	0.01	0.03**	−0.02**	−0.09*	0.02*	−0.01	0.27**	−0.05**
SEASON	0.03**	−0.03**	0.21**	0.16**	0.13**	0.10**	0.09**	−0.33**
GROWTH	0.01**	−0.01**	0.03**	0.04**	0.05**	0.03**	0.08**	−0.05**
RD	−0.02**	0.03**	−0.02**	0.04**	−0.06**	−0.02**	0.09	0.29**
TV	0.02**	−0.03**	0.01	0.01	0.01	0.01	0.70**	−0.04**

Variables	*FLLW*	*DOWN*	*VSALE*	*DISP*	*OPLEV*	*SEASON*	*GROWTH*	*RD*	*TV*
FE	0.05**	0.02	0.03**	−0.02**	0.03*	0.28**	0.01**	0.02**	0.02**
ABS-FE	−0.33**	−0.04**	0.16**	0.24**	0.10**	−0.19**	−0.01**	−0.03**	−0.04**
STICKY	0.00	0.01	0.00	−0.03**	−0.03*	0.26**	0.02**	−0.01	0.01
COGS-STICKY	0.02**	0.02**	0.01	−0.05**	−0.08**	0.18**	0.04**	0.04**	0.00
SGA-STICKY	−0.02**	0.01	0.00	−0.02**	0.00	0.13**	0.05**	−0.04**	0.00
M-STICKY	0.01**	0.05**	0.00	0.02**	−0.02**	0.14**	0.02**	−0.03**	0.02
MV	0.76**	−0.02	−0.07**	0.13**	0.18**	0.09**	0.08**	0.14**	0.72**

(continued on next page)

Variables	*FLLW*	*DOWN*	*VSALE*	*DISP*	*OPLEV*	*SEASON*	*GROWTH*	*RD*	*TV*
LOSS	−0.10**	−0.06**	0.25**	0.04**	−0.04**	−0.33**	−0.04**	0.15**	−0.05**
FLLW		−0.07**	−0.08**	−0.01	−0.10**	0.04**	0.03**	0.10**	0.61**
DOWN	0.02		0.04**	0.01	0.02	0.00	0.00	0.05**	0.01
VSALE	−0.14**	0.03**		0.03**	−0.04*	−0.03**	0.33**	0.09**	0.00
DISP	0.34**	0.01	0.02**		0.02**	−0.15**	−0.01**	0.01**	0.00
OPLEV	−0.07**	0.00	0.00	0.04**		−0.05**	0.05**	0.08**	0.01
SEASON	0.05**	0.00	−0.05**	−0.02**	−0.01**		0.12**	−0.06**	0.01
GROWTH	0.04**	0.00	0.34**	−0.01**	0.04**	0.12**		0.09**	0.10**
RD	0.08**	0.06**	0.19**	0.00	0.55**	−0.09**	0.12**		0.13**
TV	0.58**	0.00	0.01	0.01	−0.02	0.00	0.06**	0.18**	

**,* Significant at the 5 percent and 10 percent level, respectively.

Spearman coefficients are reported above the diagonal line and Pearson coefficients below the diagonal line.

Variable definitions are in Table 1.

lation with the estimated λ_2. In addition, I also estimate the correlation with industry-level coefficient estimates reported by Anderson and Lanen (2007, Table 6, Panel B). I note that Anderson et al. (2003) and Anderson and Lanen (2007) use a larger sample comprised of firms with and without analyst coverage and employ annual data. All correlation coefficients reported in Table 3 are positive and significant, indicating consistency between the proposed cost stickiness measure and the earlier evidence on the stickiness of SGA costs.

IV. RESULTS

H1 Results

To test whether stickier cost behavior results in less accurate analysts' earnings forecasts, Table 4 presents the mean and median absolute analysts' earnings forecast errors contingent on sticky ($STICKY < 0$) versus anti-sticky ($STICKY \geq 0$) cost classification. The mean absolute error for firms with sticky cost behavior is 0.0080, whereas that for firms with anti-sticky cost behavior is 0.0060. Thus, forecasts for firms with anti-sticky cost behavior are, on average, more accurate by 25 percent = (0.0080 − 0.0060)/0.0080 than forecasts for firms with sticky cost behavior. The difference is statistically significant ($p < 0.05$). If accurate earnings forecasts are valuable for capital market participants, then the difference is economically meaningful.

Table 5 presents coefficient estimates for the regression models. The coefficient on *STICKY* in model 1(a) is −0.0108, and is statistically significant ($p < 0.001$).[11] The coefficient on *COGS-*

TABLE 3

Correlation Coefficients between Industry Estimates of Cost Stickiness

Variables	Mean *SGA-STICKY*$_j$	$\hat{\lambda}_{2,j}$	Anderson-Lanen Coefficient
Mean *SGA-STICKY*$_j$		0.562**	0.345**
$\hat{\lambda}_{2,j}$	0.485**		0.467**
Anderson-Lanen Coefficient	0.463**	0.365**	

** Significant at the 5 percent level.

The table presents Spearman (Pearson) coefficients above (below) the diagonal line between three estimates of cost stickiness measured at the two-digit SIC code level: *SGA-STICKY*, an estimate based on a measure suggested by Anderson et al. (2003) and estimates reported by Anderson and Lanen (2007).

Mean *SGA-STICKY*$_j$ is the mean value of *SGA-STICKY* across all sample observations at the two-digit SIC code level, j = 20 to 39.

$\hat{\lambda}_{2,j}$ is the coefficient estimate from estimating the regression of the following model using all sample observations at the two-digit SIC code level, j = 20 to 39.

Model 4:

$$\log\left[\frac{SGA_{i,t}}{SGA_{i,t-1}}\right] = \lambda_0 + \lambda_1 \log\left[\frac{SALE_{i,t}}{SALE_{i,t-1}}\right] + \lambda_2\, SALEDEC_{i,t} \log\left[\frac{SALE_{i,t}}{SALE_{i,t-1}}\right] + \varepsilon_{i,t}$$

where $SALEDEC_{i,t}$ equals 1 if $SALE_{i,t} < SALE_{i,t-1}$, and 0 otherwise. $SGA_{i,t}$ is Compustat #1 and *SALE* is Compustat #2.

Anderson-Lanen Coefficients are taken for the respective two-digit SIC code industries from Anderson and Lanen (2007, Table 6, Panel B).

[11] Consistent with the perception of costs as sticky if firms incur disproportionate costs when activity levels decrease, results from an additional regression analysis (untabulated) indicate that cost stickiness boosts absolute earnings forecast errors more when activity levels decrease than when they increase.

TABLE 4

Absolute Forecast Errors (*ABS-FE*) for Firms with Sticky versus Anti-Sticky Cost Behavior

Cost Behavior	Mean	Median	n
Sticky Costs: $STICKY_{it} < 0$	0.0080	0.0012	23,915
Anti-Sticky Costs: $STICKY_{it} \geq 0$	0.0060	0.0010	21,016
Difference	0.0020**	0.0002[a]	

** Significant at the 5 percent level.

[a] Mann-Whitney test indicates a significant difference between the medians at the 5 percent level.

STICKY in model 1(b) is −0.0100, and is statistically significant ($p < 0.001$). The coefficient on *SGA-STICKY* in model 1(c) is −0.0055, statistically significant at $p < 0.002$. Adjusted R^2 values for the regressions vary from 7.5 percent to 17.6 percent. The results support H1, indicating that stickier cost behavior is associated with lower accuracy of analysts' earnings forecasts.

As for the control variables, results for *MV* and *LOSS* are generally consistent with expectations, indicating a positive and significant relationship between the amount of available firm-specific information and forecast error. The coefficient estimate on *DOWN* is insignificant across the regression models, possibly due to differences among analysts in the underlying costs, earnings models, and access to management information: a large number of analysts covering a firm can proxy variation in the underlying costs and profits models, resulting in considerable noise. As expected, the findings for *DISP* and to a limited extent for *VSALE* indicate a positive and significant association between the absolute magnitude of the forecast errors and the uncertainty in the firm's environment of operations and earnings predictability.

OPLEV is positively associated with *ABS-FE*, indicating that operating leverage increases the analysts' earnings forecast errors. The seasonal effect, *SEASON*, is insignificant across the regression models, indicating that analysts recognize the seasonal effect and adjust their forecasts accordingly.

Results for two sensitivity models 1(d) and 1(e) are also reported in Table 5 and provide further insights into additional aspects of the relationship between cost behavior and the accuracy of analyst earnings forecasts. First, I examine the sensitivity of the results to estimating cost stickiness over a longer time period. Accordingly, *M-ŚTICKY* measures cost stickiness based on cost responses over eight quarters. Regression results for model 1(d) indicate a statistically significant negative coefficient on *M-STICKY*, −0.0073 ($p = 0.019$). The result supports H1.

Second, I examine whether past (rather than current) managerial discretion affects the hypothesized relationship. I check whether the regression coefficient estimates are sensitive to discretionary choices made by managers in quarter $t-1$ or earlier by replacing $STICKY_{it}$ in model 1(a) with the cost stickiness measure estimated on quarter $t-1$, $STICKY_{i,t-1}$, which excludes all managerial choices made in quarter t.

Estimating regression model 1(e), the coefficient estimate on $STICKY_{i,t-1}$ is −0.0040 ($p = 0.030$). The negative and significant coefficient estimate indicates that stickier cost behavior observed in a preceding quarter is associated with higher absolute analysts' forecast errors. I conclude that cost stickiness estimated by analysts on a preceding quarter affects the accuracy of the earnings prediction.

Additionally, I examine the incremental effect of *STICKY* over earnings volatility, which is likely to be an all-inclusive noisy variable that incorporates many types of uncertainties (e.g.,

TABLE 5

Regression Coefficients of Analysts' *Absolute* Forecast Error on Cost Stickiness, Control Variables, and Two-Digit SIC-Code Industry Effects

Model 1(a)

$$ABS\text{-}FE_{it} = \beta_0 + \beta_1 STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it}$$

Model 1(b)

$$ABS\text{-}FE_{i,t} = \beta_0 + \beta_1 COGS\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it}$$

Model 1(c)

$$ABS\text{-}FE_{i,t} = \beta_0 + \beta_1 SGA\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it}$$

Model 1(d)

$$ABS\text{-}FE_{i,t} = \beta_0 + \beta_1 M\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it}$$

Model 1(e)

$$ABS\text{-}FE_{i,t} = \beta_0 + \beta_1 STICKY_{i,t-1} + \beta_2 MV_{it} + \beta_3 LOSS_{it} + \beta_4 DOWN_{it} + \beta_5 VSALE_{it} + \beta_6 DISP_{it} + \beta_7 OPLEV_{it} + \beta_8 SEASON_{it} + \varepsilon_{it}$$

Variables	Predicted Sign	Model 1(a)	Model 1(b)	Model 1(c)	Model 1(d)	Model 1(e)
Intercept	?	0.0025	0.0035	−0.0062	0.0047	0.0007
		(0.611)	(0.666)	(0.333)	(0.110)	(0.877)
$STICKY_{it}$	−	−0.0108				
		(<0.001)				
$COGS\text{-}STICKY_{it}$	−		−0.0100			
			(<0.001)			

(continued on next page)

TABLE 5 (continued)

Variables	Predicted Sign	Model 1(a)	Model 1(b)	Model 1(c)	Model 1(d)	Model 1(e)
$SGA\text{-}STICKY_{it}$	−			−0.0055		
				(0.002)		
$M\text{-}STICKY_{it}$	−				−0.0073	
					(0.019)	
$STICKY_{i,t-1}$	−					−0.0040
						(0.030)
MV_{it}	−	−0.0008	−0.0007	−0.0006	−0.0005	−0.0008
		(0.030)	(0.044)	(0.048)	(0.050)	(0.006)
$LOSS_{it}$	+	0.0082	0.0122	0.0091	0.0240	0.0095
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
$DOWN_{it}$	−	−0.0008	−0.0009	−0.0025	−0.0006	−0.0013
		(0.174)	(0.133)	(0.177)	(0.101)	(0.085)
$VSALE_{it}$	+	0.0013	0.0013	0.0022	0.0030	0.0010
		(0.041)	(0.055)	(0.080)	(0.040)	(0.065)
$DISP_{it}$	+	2.2311	1.8001	4.1311	1.6001	2.2682
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
$OPLEV_{it}$	+	0.0090	0.0082	0.0190	0.0090	0.0080
		(0.040)	(0.040)	(0.033)	(0.023)	(0.048)
$SEASON_{it}$	+	0.0001	−0.0001	0.0009	−0.0002	−0.0004
		(0.952)	(0.900)	(0.521)	(0.500)	(0.777)
n		32,563	27,411	16,918	27,811	27,401
Adj. R^2		0.176	0.122	0.075	0.136	0.186

p-values based on two-tailed tests are in parentheses.
Variable definitions are in Table 1.

demand uncertainty and operating leverage). Results (not tabulated) indicate that *STICKY* has an incremental effect on the accuracy of analysts' earnings forecasts above and beyond earnings volatility and the control variables. Overall, the evidence supports the hypothesis.[12]

Furthermore, evidence on the mean forecast error (as opposed to the absolute forecast error) reported in Table 6 offers insights into the validity of the assumption that analysts recognize cost behavior. Results show that the mean forecast error of firms with sticky costs is insignificantly different from that of firms with anti-sticky costs.[13] Thus, the evidence supports the assumption that analysts have at least some understanding of firms' cost behavior.

A final important consideration is that an analyst does not have the ability to reduce forecast errors caused by a dispersion of a firm's *ex ante* earnings distribution. In other words, an analyst cannot influence accuracy implied by cost stickiness because it is a firm-specific feature.[14] Therefore, an analyst cannot reduce the dispersion of the *ex ante* earnings distribution implied by cost stickiness even if she is aware of it in advance.

H2 Results

Results showing that firms with stickier cost behavior have lower analyst coverage are presented in Table 7. Findings in Panel A indicate that, on average, 5.459 analysts follow firms with sticky cost behavior while 5.622 analysts follow firms with anti-sticky cost behavior. The difference of about 3 percent is statistically significant ($p < 0.05$). Panel B reports the results of three regression models, 2(a), 2(b), and 2(c). The coefficients on $STICKY_{it}$, $M\text{-}STICKY_{it}$ and

TABLE 6

Forecast Errors (*FE*) for Firms with Sticky versus Anti-Sticky Cost Behavior

Cost Behavior	Mean	Median	n
Sticky Costs: $STICKY_{it} < 0$	−0.0016	0	23,915
Anti-Sticky Costs: $STICKY_{it} \geq 0$	−0.0012	0	21,016
Difference	−0.0004[a]	0	

[a] Insignificant at the 10 percent level.

[12] Results of further analyses also support H1. First, findings from estimating model 1 with a differential slope coefficient on negative stickiness (i.e., sticky costs) indicate a minor and marginally significant difference between the coefficients of negative and positive values of *STICKY* on *ABS-FE*. Second, checking for a potential seasonality effect, I also computed the stickiness measure using cost responses relative to the same quarter of the preceding year. These findings support H1.

[13] To see the intuition, suppose, on the contrary, that an analyst ignores cost stickiness. Consequently, her forecast will be upward biased in case of sticky costs (forecast error = reported earnings − forecast < 0) because she *under-estimates* costs on demand falls. In a similar vein, her forecast will be downward biased in case of anti-sticky costs (forecast error = reported earnings − forecast > 0) because she *over-estimates* costs on demand falls. Thus, sticky costs trigger a negative mean forecast error and anti-sticky costs trigger a positive mean forecast error (i.e , bias, not absolute forecast error).

However, results reported in Table 6 indicate that the mean forecast error is *not* significantly different for observations with sticky versus anti-sticky costs. Therefore, the data support the assumption that analysts recognize cost stickiness to some extent.

[14] Lys and Soo (1995) demonstrate that the inherent difficulty in predicting earnings is associated with large forecast errors (see also Kross et al. 1990). Alford and Berger (1999, 219) suggest a proxy for "analysts' *ability* to predict company's earnings" (emphasis added). In contrast, firm-specific sticky costs increase the *ex ante* dispersion of the firm's earnings distribution. The correlation between *STICKY* (*M-STICKY*) and this proxy (using Equation (1) in Alford and Berger 1999, 223) is −0.07 (0.04), suggesting that the two variables do not pick up the same phenomena.

TABLE 7

Association of Cost Behavior with Analyst Coverage

Panel A: Mean Number of Analysts Following Firms with Sticky versus Anti-Sticky Cost Behavior.

Cost Behavior	Mean Number of Analyst Coverage	n
Sticky Costs: $STICKY_{it} < 0$	5.459	23,915
Anti-Sticky Costs: $STICKY_{it} \geq 0$	5.622	21,016
Difference	0.163**	

Panel B: Regression of the Number of Analysts Following Firms on Cost Stickiness, Control Variables, Year Effects and Two-Digit SIC-Code Industry Effects

Model 2(a)

$$FLLW_{it} = \beta_0 + \beta_1 STICKY_{it} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it}$$

Model 2(b)

$$FLLW_{it} = \beta_0 + \beta_1 M\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it}$$

Model 2(c)

$$FLLW_{it} = \beta_0 + \beta_1 STICKY_{i,t-1} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it}$$

Variables	Model 2(a)	Model 2(b)	Model 2(c)
Intercept	0.0500	0.0448	0.0586
	(<0.001)	(<0.001)	(<0.001)
$STICKY_{it}$	0.0211		
	(0.031)		
$M\text{-}STICKY_{it}$		0.0144	
		(0.042)	
$STICKY_{i,t-1}$			0.0216
			(0.033)
MV_{it}	0.3174	0.3188	0.3171
	(<0.001)	(<0.001)	(<0.001)
RD_{it}	0.2196	0.2977	0.2633
	(<0.001)	(<0.001)	(<0.001)
$VSALE_{it}$	−0.4133	−0.5001	−0.4702
	(<0.001)	(<0.001)	(<0.001)
$DISP_{it}$	0.6735	0.6448	0.7116
	(<0.001)	(<0.001)	(<0.001)
$ABS\text{-}FE_{it}$	−0.0854	−0.0998	−0.0796
	(0.023)	(0.046)	(0.045)
$GROWTH_{it}$	−0.0067	−0.0444	0.0360
	(0.786)	(0.171)	(0.219)
TV_{it}	0.1551	0.1881	0.1776
	(<0.001)	(<0.001)	(<0.001)

(continued on next page)

Panel B: Regression of the Number of Analysts Following Firms on Cost Stickiness, Control Variables, Year Effects and Two-Digit SIC-Code Industry Effects

Model 2(a)

$$FLLW_{it} = \beta_0 + \beta_1 STICKY_{it} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it}$$

Model 2(b)

$$FLLW_{it} = \beta_0 + \beta_1 M\text{-}STICKY_{it} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it}$$

Model 2(c)

$$FLLW_{it} = \beta_0 + \beta_1 STICKY_{i,t-1} + \beta_2 MV_{it} + \beta_3 RD_{it} + \beta_4 VSALE_{it} + \beta_5 DISP_{it} + \beta_6 ABS\text{-}FE_{it} + \beta_7 GROWTH_{it} + \beta_8 TV_{it} + \beta_9 LOSS_{it} + \varepsilon_{it}$$

Variables	Model 2(a)	Model 2(b)	Model 2(c)
$LOSS_{it}$	−0.0059	−0.0055	−0.0023
	(0.094)	(0.111)	(0.188)
n	35,857	31,532	31,662
Pseudo-R^2	43.18%	45.81%	44.07%

** Significant at the 5 percent level.

The regression model was estimated using a standard negative binomial distribution because the dependent variable (*FLLW*) is count-data. The dispersion parameter was estimated by maximum likelihood. p-values are reported in parentheses. The pseudo-R^2, also named McFadden's R^2, is the log-likelihood value on a scale from 0 to 1, where 0 corresponds to the constant-only model and 1 corresponds to perfect prediction (a log-likelihood of 0).

Variable definitions are in Table 1.

$STICKY_{i,t-1}$ are positive and highly significant. Keeping in mind that lower values of *STICKY* indicate stickier cost behavior, the findings reject the null H2.[15]

As for the control variables, the coefficient estimates of *MV* and *TV* are positive and significant, in line with prior research. The coefficient estimates of the proxies for environmental uncertainty show mixed results. The coefficients of *VSALE* and *ABS-FE* are negatively and significantly associated with the analyst following, while the coefficient of *DISP* is positive and significant. The coefficients of *GROWTH* and *LOSS* are insignificant.

The coefficient estimate of *RD* is also positive and highly significant, consistent with Barth et al. (2001). To further check the robustness of the cost behavior effect, I separately examine the cost stickiness effect on analyst coverage for firms with and without R&D expenditures. Results (not tabulated) indicate that cost stickiness is significantly associated with analyst coverage for firms with and without R&D expenditures. In sum, the evidence indicates that firms with stickier cost behavior have lower analyst coverage.

Lower coverage for firms with stickier costs and more volatile earnings may seem counterintuitive if analysts strive to meet a high demand for earnings forecasts for firms that have less predictable earnings. However, the analysts' attitude toward large negative forecast errors can

[15] The analysis implicitly assumes that an equivalent effort is expended for estimating sticky and anti-sticky costs. This assumption is sensible in this context because cost stickiness is estimated from public information reported in financial statements.

partially explain their coverage preferences. Ample evidence shows substantial declines in share price following a negative forecast error (e.g., Bartov et al. 2002).[16] To some extent, analysts' short- and long-term benefits are affected by their relationships with managers of covered firms (Lim 2001). Therefore, all else being equal, analysts are likely to prefer covering firms with low *ex ante* probability of large negative forecast errors. Risk aversion reflected in a conventional concave loss-utility function captures these preferences. This interpretation implicitly assumes some disparity in risk attitude to large negative forecast errors between investors and analysts or, alternatively, that investors recognize cost stickiness to a limited extent.

H3 Results

Table 8 presents results from testing whether the market response to earnings surprises is weaker for firms with stickier cost behavior. In line with the prior literature, coefficient estimates β_1 in all regression models are positive and highly significant, indicating a positive market response to earnings surprises. The estimated coefficients for the interaction variable are positive and significant when cost stickiness relates to total costs (models 3(a) and 3(d)), but only marginally significant when cost stickiness relates to *SGA* costs (model 3(c)), and insignificant with respect to stickiness of *COGS* (model 3(b)). Additionally, results from estimating model 3(e), which uses an indicator variable for the classification of costs as sticky versus anti-sticky, support H3.

The findings suggest that investors recognize and consider cost stickiness with respect to total costs, but not the stickiness of cost components. The explanatory power in the models ranges between 1.8 percent and 2.9 percent, which is in line with prior literature (e.g., Gu and Wu 2003). To strengthen the evidence, I take a predictive rather than contemporaneous approach in estimating cost stickiness. Model 3(f) shows a lower market reaction to earnings surprises for firms with less sticky costs estimated on the preceding quarter (note that $STICKY < 0$ indicates sticky costs). Taken as a whole, the findings corroborate Banker and Chen (2006) and indicate a weaker market response to earnings surprises for firms with stickier cost behavior, supporting H3.

These results contribute to the ongoing debate on investor rationality by documenting that investors are able to process accounting information and partially infer cost behavior in a rational manner. With respect to the control variables, coefficient estimates for *DISP* are generally insignificant and coefficient estimates for *VSALE* are only marginally significant. Thus, dispersion of analysts' forecasts and variation of sales may not serve as appropriate proxies for *ex ante* earnings predictability as perceived by investors. This argument is supported by Diether et al. (2002), who interpret dispersion in analysts' earnings forecasts as a proxy for differences in opinion about the stock (e.g., due to the employment of different valuation models). While forecast dispersion may indicate different opinions or the use of different forecasting models, cost stickiness serves as a proxy for more volatile earnings due to firm-specific cost structures. Thus, the two proxies capture different aspects of earnings predictability.[17] Overall, findings indicate that investors have at least some understanding of firms' cost behavior in responding to earnings surprises.

V. A CONCLUDING REMARK

The study utilizes a managerial accounting concept, sticky costs, to gain insights into how firms' cost behavior affects (1) the accuracy of analysts' earnings forecasts, (2) analysts' selection of covered firms, and (3) the market response to earnings announcements. While implications of cost behavior are of primary interest to management accountants, this study employs a manage-

[16] Kinney et al. (2002) provide a different view, which finds considerable variation in returns for firms reporting positive or negative surprises.

[17] See Dichev and Tang (2009) and Frankel and Litov (2009) for additional aspects of earnings predictability.

TABLE 8

Effect of Sticky Cost Behavior on Stock Market's Reaction to Earnings Surprises

Model 3(a)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it}$$

Model 3(b)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} COGS\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it}$$

Model 3(c)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} SGA\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it}$$

Model 3(d)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} M\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it}$$

Model 3(e)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} D\text{-}STICKY_{it} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it}$$

Model 3(f)

$$CAR_{it} = \beta_0 + \beta_1 FE_{it} + \beta_2 FE_{it} STICKY_{i,t-1} + \beta_3 DISP_{it} + \beta_4 VSALE_{it} + \varepsilon_{it}$$

Variable	Model 3(a)	Model 3(b)	Model 3(c)	Model 3(d)	Model 3(e)	Model 3(f)
Intercept	0.0208	0.0197	0.0256	0.0355	0.0011	0.0244
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.066)	(<0.001)
FE_{it}	0.2929	0.3326	0.3636	0.3467	0.4377	0.3745
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
$FE_{it}STICKY_{it}$	0.0166					
	(0.033)					
$FE_{it}COGS\text{-}STICKY_{it}$		0.0238				
		(0.141)				
$FE_{it}SGA\text{-}STICKY_{it}$			0.0089			
			(0.077)			
$FE_{it}M\text{-}STICKY_{it}$				0.0202		
				(0.038)		
$FE_{it}D\text{-}STICKY_{it}$					0.0366	
					(0.040)	
$FE_{it}STICKY_{i,t-1}$						0.0244
						(0.048)

(continued on next page)

TABLE 8 (continued)

Variable	Model 3(a)	Model 3(b)	Model 3(c)	Model 3(d)	Model 3(e)	Model 3(f)
$DISP_{it}$	0.9202	0.8333	1.0284	1.0196	1.1280	1.1017
	(0.151)	(0.095)	(0.156)	(0.640)	(0.251)	(0.233)
$VSALE_{it}$	0.0120	0.0377	0.0095	0.0111	0.0267	0.0212
	(0.048)	(0.089)	(0.090)	(0.050)	(0.055)	(0.075)
Adj. R^2	0.029	0.018	0.019	0.027	0.028	0.019
n	43,777	37,389	23,666	38,121	43,777	37,119

p-values based on two-tailed tests are in parentheses.

Variable Definitions:

CAR_{it} = cumulative market-adjusted returns (raw return minus value-weighted CRSP return) measured over three trading days surrounding earnings announcement, from the day before to the day after; and

$D\text{-}STICKY_{it} = -1$ if $STICKY_{it} < 0$, and 0 otherwise.

Definitions of the other variables are in Table 1.

ment accounting concept for addressing research questions usually raised by financial accountants. Although such a multi-disciplinary approach has not been common in the prior literature, the findings of this study suggest that combining the perspectives of management and financial accounting can be fruitful. Further research is expected to build on this approach in exploring multi-disciplinary accountings topics. Integrating management and financial accounting research is likely to benefit both disciplines.

APPENDIX

Employing cost stickiness as a yardstick, I develop a simple two-period model to predict a relationship between the level of cost stickiness and the accuracy of analysts' earnings forecasts. The primitive model input in the first period is a set of prior beliefs on the state of the world, say demand y, which is a realization of a random variable $\tilde{y}$, drawn from a distribution function, $\Phi(y)$, with a strictly positive and symmetric density, $\phi(y)$, over the support $[\bar{y} - \lambda, \bar{y} + \lambda]$, $\bar{y} > \lambda > 0$. The second-period revenue function, $R(y)$, is assumed to be differentiable, increasing and concave.[18] Costs in the second period are modeled by:

$$C(y, \bar{y}, \alpha) = \begin{cases} f + (v + \alpha)y & \text{if } y < \bar{y} \\ f + vy + \alpha\bar{y} & \text{if } y \geq \bar{y} \end{cases}$$

f = a fixed cost of production, $f \geq 0$;
v = a variable cost per product unit, $v \geq 0$; and

[18] The revenue function depends on previously made managerial choices, like product price. See also Balakrishnan and Sivaramakrishnan (2002).

α = level of cost stickiness per product unit, $-v < \alpha < R'(\bar{y}) - v$.

The parameter α captures the level of cost stickiness in adjusting resources to changes in the activity level. If $\alpha < 0$, then costs increase *more* when activity rises than they decrease when activity falls by an equivalent amount; that is, a negative value of α indicates sticky costs. If $\alpha > 0$, then costs increase *less* when activity rises than they decrease when activity falls by an equivalent amount; that is, a positive value of α indicates anti-sticky costs. The difference between the cost of an upward activity adjustment and the cost of an equivalent downward activity adjustment depends only on α:[19]

$$[C(\bar{y}+1) - C(\bar{y})] - [C(\bar{y}) - C(\bar{y}-1)] = [v(\bar{y}+1) + \alpha\bar{y} - v\bar{y} - \alpha\bar{y}] - [v\bar{y} + \alpha\bar{y} - (v+\alpha)(\bar{y}-1)]$$
$$= -\alpha.$$

I use cost of adjustments to expand the conventional fixed-variable cost model and estimate stickiness of firms' cost functions. My approach follows Wernerfelt (1997), who shows that the magnitude of an adjustment cost drives the form of the organization, and Balakrishnan and Gruca (2008), who show that cost stickiness is greater for cost functions that relate to an organization's core competency. Rothschild (1971) models properties of convex (concave) adjustment cost structures that result in asymmetric cost functions due to the cost of producing marginal unit increases (decreases) in the activity level.

In my model, the earnings function, $\Pi(y,\alpha) = R(y) - C(y,\alpha)$, is strictly increasing in y and transforms demand y realized in the second period into earnings. The *ex ante* earnings expectations in the first period are denoted $\hat{\Pi}(\tilde{y}, \alpha)$. In the second period, the firm truly reports its realized earnings, $\Pi(y,\alpha)$.

An analyst is delegated the task of producing accurate estimates of a firm's earnings expectations and the forecast is honest, as in, for example, Ottaviani and Sorensen (2006). In the first period, the analyst announces $\hat{\Pi}(\tilde{y}, \alpha)$ as her most accurate forecast, if her error loss function is symmetric and concave (e.g., a quadratic loss function).[20] Focusing on the absolute earnings forecast error as an accuracy gauge, the proposition below proves a negative relationship between the level of cost stickiness and the mean absolute earnings forecast error. That is, higher values of α, i.e., less sticky cost behavior, result in lower mean absolute analyst forecast errors.

Proposition

$E[|\Pi(\tilde{y}, \alpha) - \hat{\Pi}(\tilde{y}, \alpha)|]$ decreases in α.

Proof

The proof is based on Jensen's inequality:

$$\text{Let } FE(y,\alpha) = \Pi(y,\alpha) - \hat{\Pi}(y,\alpha) = \begin{cases} R(y) - f - (v+\alpha)y - \hat{\Pi}(\tilde{y},\alpha) & \text{if } y < \bar{y}, \\ R(y) - f - vy - \alpha\bar{y} - \hat{\Pi}(\tilde{y},\alpha) & \text{otherwise.} \end{cases}$$

Define $y^{\#}(\alpha)$: $\Pi(y^{\#}(\alpha), \alpha) = \hat{\Pi}(\tilde{y}, \alpha)$. Thus, $FE(y, \alpha) \leq 0$ for all $y \in [\bar{y} - \lambda, y^{\#}(\alpha)]$.

[19] I note that cost stickiness does not depend on the operating leverage of the firm because the fixed cost component, f, does not influence the level of cost stickiness.

[20] A discussion on the properties of a symmetric error loss function appears in Beja and Weiss (2006).

$$E[|FE(\tilde{y},\alpha)|] = \int_{\bar{y}-\lambda}^{y^{\#}(\alpha)} -\mathrm{FE}(y,\alpha)\phi(y)\mathrm{d}y + \int_{y^{\#}(\alpha)}^{\bar{y}+\lambda} FE(y,\alpha)\phi(y)\mathrm{d}y = -2\int_{\bar{y}-\lambda}^{y^{\#}(\alpha)} FE(y,\alpha)\phi(y)\mathrm{d}y$$

$$+\int_{\bar{y}-\lambda}^{\bar{y}+\lambda} FE(y,\alpha)\phi(y)\mathrm{d}y = -2\int_{\bar{y}-\lambda}^{y^{\#}(\alpha)} FE(y,\alpha)\phi(y)\mathrm{d}y > 0$$

Jensen's inequality implies $\alpha_1 \geq \alpha_2 \Rightarrow y^{\#}(\alpha_1) \geq y^{\#}(\alpha_2)$, and, $FE(y, \alpha_2) \leq \mathrm{FE}(y, \alpha_1)$ for all $y \in [\bar{y} - \lambda, y^{\#}(\alpha_2)]$.

Suppose $\alpha_1 \geq \alpha_2$. Hence:

$$E[|FE(\tilde{y},\alpha_2)|] - E[|FE(\tilde{y},\alpha_1)|] = -2\int_{\bar{y}-\lambda}^{y^{\#}(\alpha_2)} FE(y,\alpha_2)\phi(y)\mathrm{d}y + 2\int_{\bar{y}-\lambda}^{y^{\#}(\alpha_1)} FE(y,\alpha_1)\phi(y)\mathrm{d}y$$

$$= 2\left[\int_{\bar{y}-\lambda}^{y^{\#}(\alpha_2)} (FE(y,\alpha_1) - FE(y,\alpha_2))\phi(y)\mathrm{d}y\right]$$

$$-2\int_{y^{\#}(\alpha_2)}^{y^{\#}(\alpha_1)} FE(y,\alpha_1)\phi(y)\mathrm{d}y \geq 0.$$

Based on the proposition, I argue that cost stickiness is a determinant of the accuracy of analysts' earnings forecasts. Specifically, more sticky costs increase the spread of the *ex ante* distribution of earnings, which increases the *ex ante* volatility of reported earnings. For that reason the proposition motivates H1: Increased cost stickiness reduces the accuracy of analysts' consensus earnings forecasts.

REFERENCES

Abarbanell, J. S., and B. J. Bushee. 1997. Fundamental analysis, future earnings, and stock returns. *Journal of Accounting Research* 35 (1): 1–24.

———, W. Lanen, and R. Verrecchia. 1995. Analysts' forecasts as proxies for investor beliefs in empirical research. *Journal of Accounting and Economics* 20: 31–60.

Adar, Z., A. Barnea, and B. Lev. 1977. A comprehensive cost-volume-profit analysis under uncertainty. *The Accounting Review* 52: 137–149.

Alford, W. A., and P. G. Berger. 1999. A simultaneous equations analysis of forecast accuracy, analyst following, and trading volume. *Journal of Accounting, Auditing & Finance* 14: 219–240.

Anderson, M., R. Banker, and S. Janakiraman. 2003. Are selling, general, and administrative costs 'sticky'? *Journal of Accounting Research* 41: 47–63.

———, ———, R. Huang, and S. Janakiraman. 2007. Cost behavior and fundamental analysis of SGA costs. *Journal of Accounting, Auditing & Finance* 22 (1): 1–22.

Anderson, S., and W. Lanen. 2007. Understanding cost management: What can we learn from the evidence on "sticky costs"? Working paper, University of Melbourne.

Atiase, R. 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research* 23: 21–36.

Baik, B. 2006. Self-selection in consensus analysts' earnings forecasts. *Asia-Pacific Journal of Financial*

Studies 35 (6): 141–168.

Balakrishnan, R., and K. Sivaramakrishnan. 2002. A critical overview of the use of full-cost data for planning and pricing. *Journal of Management Accounting Research* 14: 3–31.

———, M. J. Peterson, and N. S. Soderstrom. 2004. Does capacity utilization affect the "stickiness" of costs? *Journal of Accounting, Auditing & Finance* 19 (3): 283–299.

———, and T. S. Gruca. 2008. Cost stickiness and core competency: A note. *Contemporary Accounting Research* 25 (4): 993–1006.

———, and N. S. Soderstrom. 2009. Cross-sectional variation in cost stickiness. Working paper, The University of Iowa.

Banker, R., and L. Chen. 2006. Predicting earnings using a model based on cost variability and cost stickiness. *The Accounting Review* 81: 285–307.

———, M. Ciftci, and R. Mashruwala. 2008. Managerial optimism, prior-period sales changes, and sticky cost behavior. Working paper, Temple University.

Barron, O. E., O. Kim, S. Lim, and D. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73: 421–433.

———, D. Byard, C. Kile, and E. J. Riedl. 2002. High-technology intangibles and analysts' forecasts. *Journal of Accounting Research* 40 (2): 289–312.

Barth, M. E., R. Kasznik, and M. F. McNichols. 2001. Analyst coverage and intangible assets. *Journal of Accounting Research* 39 (1): 1–34.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33: 173–204.

Beja, A., and D. Weiss. 2006. Some informational aspects of conservatism. *European Accounting Review* 15: 585–604.

Bhushan, R. 1989. Firm characteristics and analyst following. *Journal of Accounting and Economics* 11: 255–274.

Brown, L. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39: 221–241.

———, G. Richardson, and S. Schwager. 1987. An information interpretation of financial analyst superiority in forecasting earnings. *Journal of Accounting Research* 25: 49–67.

Chen, C. X., H. Lu, and T. Sougiannis. 2008. Managerial empire building, corporate governance, and the asymmetrical behavior of selling, general, and administrative costs. Working paper, University of Illinois at Urbana–Champaign.

Clement, M., and S. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78: 227–249.

Collins, D. W., S. P. Kothari, and J. D. Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics* 9: 111–138.

Das, S., C. Levine, and K. Sivaramakrishnan. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73: 277–294.

———, R. J. Gou, and H. Zhang. 2006. Analysts' coverage and subsequent performance of newly public firms. *The Journal of Finance* 61 (3): 1159–1185.

Dechow, P., S. Kothari, and R. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

Dichev, I. D., and V. W. Tang. 2009. Earnings volatility and earnings predictability. *Journal of Accounting and Economics* 47: 160–181.

Diether, K., C. Malloy, and A. Scherbina. 2002. Differences of opinion and the cross section of stock returns. *The Journal of Finance* 57: 2113–2141.

Frankel, R., S. Kothari, and J. Weber. 2006. Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41: 29–54.

———, and L. Litov. 2009. Earnings persistence. *Journal of Accounting and Economics* 47: 182–190.

Givoly, D., and C. Hayn. 2000. The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? *Journal of Accounting and Economics* 29: 287–320.

Gu, Z., and J. Wu. 2003. Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics* 35: 5–29.

Hemmer, T., and E. Labro. 2008. On the optimal relation between the properties of managerial and financial reporting. *Journal of Accounting Research* 46 (5): 1209–1240.

Hong, H., T. Lim, and J. C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance* 55: 265–295.

Imhoff, E., and G. Lobo. 1992. The effect of *ex ante* earnings uncertainty on earnings response coefficients. *The Accounting Review* 67: 427–439.

Kinney, W., D. Burgstahler, and R. Martin. 2002. Earnings surprise "materiality" as measured by stock returns. *Journal of Accounting Research* 40 (5): 1297–1329.

Kross, W., B. Ro, and D. Schroeder. 1990. Earnings expectations: The analysts' information advantage. *The Accounting Review* 65 (2): 461–476.

Lang, M., and R. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71: 467–492.

Lev, B., and R. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* 31: 190–215.

Lim, T. 2001. Rationality and analysts' forecast bias. *The Journal of Finance* 56: 369–385.

Lipe, R. 1990. The relation between stock returns and accounting earnings given alternative information. *The Accounting Review* 65: 49–71.

Lys, T., and L. G. Soo. 1995. Analysts' forecast precision as a response to competition. *Journal of Accounting, Auditing & Finance* 10 (4): 751–766.

Maher, M., W. Lanen, and M. Rajan. 2006. *Fundamentals of Cost Accounting*. New York, NY: McGraw-Hill.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77: 483–514.

O'Brien, P. C., and R. Bhushan. 1990. Analyst following and institutional ownership. *Journal of Accounting Research* 28: 55–82.

Ottaviani, M., and P. Sorensen. 2006. The strategy of professional forecasting. *Journal of Financial Economics* 81: 441–466.

Parkash, M., D. S. Dhaliwal, and W. K. Salatka. 1995. How certain firm-specific characteristics affect the accuracy and dispersion of analysts' forecasts. *Journal of Business Research* 34: 161–169.

Philbrick, D., and W. Ricks. 1991. Using value line and IBES analyst forecasts in accounting research. *Journal of Accounting Research* 29: 397–471.

Rock, S., S. Sedo, and M. M. Willenborg. 2000. Analyst following and count-data econometrics. *Journal of Accounting and Economics* 30: 351–373.

Rothschild, M. 1971. On the cost of adjustment. *The Quarterly Journal of Economics* 85: 601–622.

Stickel, S. 1992. Reputation and performance among security analysts. *The Journal of Finance* 47: 1811–1836.

Weiss, D., P. A. Naik, and C. L. Tsai. 2008. Extracting forward-looking information from security prices: A new approach. *The Accounting Review* 83 (4): 1101–1124.

Wernerfelt, B. 1997. On the nature and scope of the firm: An adjustment-cost theory. *The Journal of Business* 80 (4): 489–514.

Wiedman, C. 1996. The relevance of characteristics of the information environment in the selection of a proxy for the market's expectations for earnings: An extension of Brown, Richardson, and Schwager. *Journal of Accounting Research* 34: 313–324.

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1473–1474

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1473

BOOK REVIEWS

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Business, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

ROBERT BLOOMFIELD and KRISTINA RENNEKAMP, *Experimental Research in Financial Reporting: From the Laboratory to the Virtual World*, Foundations and Trends® in Accounting (Boston, MA: now Publishers, Inc., 2009, ISBN 978-1-60198-214-8, vol. 3, no. 1, pp. ix, 89).

In this monograph, Robert Bloomfield and Kristina Rennekamp make the case that accounting researchers can benefit from using virtual worlds as a laboratory to study the impacts of complex institutions on aspects of the financial reporting process. After a brief introductory chapter, the monograph proceeds with a description of research in financial accounting and a discussion of the benefits of experiments for learning about behavior in financial reporting contexts, and then moves in Chapter 2 to a discussion of costs and benefits of including institutional complexity in accounting experiments. Chapters 3 and 4 describe two important virtual worlds and how their economies work. Chapters 5 through 7 discuss the various ways of using virtual worlds to conduct experiments, methods for implementing some key institutions, and potential problems with conducting experiments embedded in virtual worlds and ways of mitigating these problems. Chapter 8 contains a discussion of the costs and benefits of using virtual worlds to conduct research across research methodologies. Chapter 9 contains concluding remarks.

A variety of audiences will find aspects of the monograph to be interesting and informative. As a nongamer, I was surprised to learn that Second Life has dance parties and concerts. I was even more surprised to learn that it has an active financial sector, including several stock exchanges, that there are external markets for its currency, and that its owners had to ban banking activities after allegations of fraud. These facts confirmed for me the "reality" of virtual worlds and piqued my interest in using them as a research tool.

As a researcher interested in financial reporting issues, I found the authors' discussion of the benefits of adding complexity to experiments (Chapter 2) compelling. The authors point out that adding complex institutions (defined broadly to include laws and regulations, common types of commercial and social contracts, and interactions, etc.) to experiments allows examination of a broader set of questions. They provide several examples of research questions that can be studied in this sort of experiment. I found the list, which includes examining the effects of institutional changes on the reliability of financial reports, on how information intermediaries do their work, on commercial arrangements, on real investment and production decisions, along with the authors' examples, to be interesting and thought provoking. It prompted me to consider how I could use virtual worlds to address issues of particular interest to me.

Specifically, while the authors do not explicitly make this point, their list makes clear that the set of researchable questions is expanded when complex institutions are incorporated into experiments, because interactions among individuals and between individuals and institutions are allowed. Thus, the set of possible dependent measures is expanded to include joint outcomes, and multiple levels of "play" allow examination of aggregate (e.g., economy-level) outcomes.

For example, consider controversial policy issues in accounting, such as the likely impact of adopting less

precise accounting standards on auditor liability (e.g., Schipper 2003; Benston et al. 2006). While careful thinking can identify several of the critical causal factors (e.g., Hail et al. 2009) and laboratory experimentation can identify additional factors and test theoretical relationships (Kadous and Mercer 2010), it is probably impossible to anticipate all the effects of introducing a change in institution into a complex environment, such as the U.S. economy, and it is unlikely that researchers working in this fashion can fully anticipate how various factors will play against each other. This leads to what I see as the biggest advantage of adding complexity to experiments in a virtual world environment—the opportunity for relatively low-cost, relatively efficient theory development. That is, in addition to allowing testing of hypotheses about complex institutions, virtual world experimentation can serve accounting research in the scientific role that is traditionally reserved for case studies and field experiments—that of identifying new variables and relationships for development of theory, which can then be further tested in controlled experimental settings or in archival data if the policy change is implemented. And they can do so without the cost and time lags that would occur if the institutional changes were actually implemented in the real (as opposed to virtual) world.

While the early chapters inspire the reader to rush out to virtual worlds to conduct experiments, the later chapters counteract this effect, to some extent. Specifically, the authors' discussion of the various ways of conducting studies in virtual worlds (Chapters 5 and 8) leads one to quickly conclude that the broadest research questions are out of reach unless one is able to corroborate with developers of virtual worlds. Also, there is a disconnect in the monograph in that the discussion of how some key institutional details are implemented in virtual worlds (Chapter 6) focuses on how reporting and enforcement are automatic and can be perfectly enforced in virtual worlds, which seems to rule out investigating more interesting questions that rely on discretion in reporting and enforcement. On the other hand, the discussion in Chapter 7 about difficulties of conducting experiments in virtual worlds and potentially mitigating factors is quite practical and is of immediate use to new researchers or those just beginning to conduct experiments online, whether in virtual worlds or not.

Another positive aspect of the monograph is that it discusses of the types of studies that are well-suited and ill-suited for virtual worlds. The reader gets a concentrated dose of this material in Chapter 8, but it is also sprinkled throughout the rest of monograph. Thus, the monograph is not an indiscriminate booster for the virtual worlds technology, but instead it provides a balanced discussion of the types of research questions that would and would not benefit from the virtual worlds technology.

In sum, Bloomfield and Rennekamp's monograph provides an introduction to the use of virtual worlds in accounting research. In doing so, it provides new researchers with good coverage of some basic research design issues, "old" researchers with some fresh new inspiration, and perhaps all of us with a glimpse into the future of accounting research.

REFERENCES

Benston, G. J., M. Bromwich, and A. Wagenhofer. 2006. Principles- versus rules-based accounting standards: The FASB's standard setting strategy. *Abacus* 42 (2): 165–188.

Hail, L., C. Leuz, and P. Wysocki. 2009. Accounting convergence and the potential adoption of IFRS by the United States: An analysis of economic and policy factors. Working paper, The Wharton School.

Kadous, K., and M. Mercer. 2010. Jury verdicts against auditors under precise and imprecise accounting standards. Working paper, Emory University.

Schipper, K. 2003. Principles-based accounting standards. *Accounting Horizons* 17 (1): 61–72.

KATHRYN KADOUS
Associate Professor, Goizueta Business School
Emory University

ROBERT S. KAPLAN and DAVID P. NORTON , *The Execution Premium: Linking Strategy to Operations for Competitive Advantage* (Boston, MA: Harvard Business Press, 2008, ISBN 13: 978-1-4221-2116-0, pp. xiii, 320).

If you are an admirer of Kaplan and Norton's earlier work on the balanced scorecard, then you will find much to be pleased with in their most recent book, *The Execution Premium*. If you found the balanced scorecard framework too narrow, you will also be pleased with their new book, because it contextualizes the balanced scorecard (BSC) within the larger framework of strategy development and execution. If you are a practicing manager looking for sound advice about strategy execution, you will find much that is helpful. If you are a pure academician looking for a book driven by empirical data that has been subjected to rigorous statistical analysis, then you will be disappointed by the nonscientifically selected sample of cases which form the body of the data for this book. However, I would still advise the purists to read the book, because they can use the practical examples and case studies to enrich their classes and discover rich research opportunities to extend our understanding of strategy and strategy execution in organizations.

For admirers of Kaplan and Norton's prior works, this book continues the tradition of providing easy to understand and execute tools and techniques for strategy development and execution. There is a simple framework that ties together work on strategy development, strategy mapping, the balanced scorecard, and activity-based costing.

The book begins by providing an overview of strategy development. This is a review of the traditional frameworks for strategy formulation, such as SWOT and PESTEL analysis. Chapter 3 provides a useful discussion about planning the strategy execution. It is a useful reminder that "planning the plan" is quite separate from the planning effort itself. Chapter 4 continues the theme of planning the strategy by discussing the strategic initiatives that need to be launched to execute the strategy. This is a crucial chapter which links strategy, operations, and budgets. It forces a manager to ask what changes and operations are necessary to execute the strategy and what resources he or she will need to drive the operational changes. Chapter 5 discusses the all-important structural and behavioral issues involved in aligning organizational units with the strategy. This is probably one of the weakest chapters in the book, because it does not deal with the messy problems of change management, leadership, and motivation. In my opinion, the authors rely too heavily on structural solutions such as cascading the BSC or creating a communications plan. In reality, alignment requires a shared frame of reference that cannot be achieved through structural mechanisms alone. Chapters 6 and 7 cover operational improvements, using tools such as traditional budgeting, activity-based costing, and capacity analysis. Most of the material in these chapters would not be new for accountants. However, they are extremely useful for managers who may not appreciate the strategic use of activity-based costing. The next two chapters deal with operational and strategic review meetings to adapt and change the strategy in the light of changed circumstances. The last chapter concludes by arguing effectively for the creation of an office of strategy management so that there is a clear owner of strategy and strategy execution.

For those of the belief that the balanced scorecard framework was too narrow, Kaplan and Norton have admirably redressed this imbalance by framing the BSC as part of a larger framework of strategy development and execution. I happen to be one of those who always felt that measurement without a sense of strategy would probably do more harm than good in an organization. My belief is based on a half century of research in the behavioral sciences that shows that when people do not understand or subscribe to the purpose of the measurement, they are likely to undertake dysfunctional behaviors. I am therefore delighted to see that the current book recognizes the importance of communicating with, and aligning, people and strategy before putting in a scorecard.

For practicing managers, the book provides a lot of advice that will be helpful as they embark on their own exercise of strategy execution. In particular, the concept of an office of strategy execution is an important and underappreciated concept. This is probably because of the bad experience most corporations had with the strategic planning office, which often became the place for unsuccessful line managers.

If you are an academician who believes in empirical data and rigorous statistical analysis, you will find very little of it in this book. Most of the data in this book comes from Harvard Business School teaching cases or from the consulting practice of Kaplan and Norton. From an empirical perspective, the flaws in the data are obvious. The sample is nonscientific; it comes mostly from opportunistic interventions. It is a bit paradoxical that a book which is selling a rational-scientific methodology for strategy development and execution uses cases as opposed to a matched or paired sample methodology to show that the group with tight linkage between strategy execution and operational improvement has better results than one that does not. Even the data for firms that have performed well with a balanced scorecard and other mechanisms for sound strategy execution must be taken with a grain of salt.

This is because the study covers a period from 2000 to 2006, one in which the stock market rose because of reasons that have come back to bite us in the year 2008. For example, consider the case of KeyCorp—one of the success stories cited in the book. All of its financial gains have been nearly wiped out, and I presume that the company still uses the balanced scorecard.

Paradoxically, for future doctoral students and researchers looking for topics in managerial accounting, this book's weaknesses are a treasure trove of research opportunities. While Kaplan and Norton have started to enlarge the context in which strategy is developed and executed, a large part of the domain of strategy execution remains unmapped. The world of Kaplan and Norton is a simple, rational, ordered place. They see the problem primarily in terms of mapping relationships between organizational structures and processes. Messy things such as organizational power and politics, behavioral and motivational issues, and organizational culture are neatly sidestepped. They are simply lumped into the "learning and growth" category which turns into the black box that holds all the misbehaving children that bedevil strategy executors in the real world.

Most practicing managers, however, will tell you that strategy execution is a very messy process, because it has to negotiate the tricky waters of organizational power and politics, organizational culture, and employee motivation. These are not only significant challenges, but they require very different tools. While the balanced scorecard and activity-based costing can help to understand weakness in performance, aligning employees requires engagement and dialog. Leadership in these situations is not just about communications that inform or tell employees, but requires engaging employees. Leadership becomes a matter of teaching and learning and not simply communicating.

A great opportunity for future research comes from critically examining the process of strategy formulation and execution from a different perspective. Kaplan and Norton's work is rooted in the rational-predictive model of managerial behavior. Their work assumes that there is a knowable but uncertain environment, and a manager's task is to identify opportunities through SWOT analysis and then exploit them through strategy execution. This may work for organizations that operate under conditions of uncertainty or which can control the uncertainty through predictive methods. But what happens when organizations face not just *uncertainty* but an *unknowable* environment? How does strategy formulation and execution proceed in such situations?

One answer to this question comes from the work of Professor Saras D. Sarasvathy of the University of Virginia, who has studied the behavior of entrepreneurs. Future scholars interested in the study of strategy should read Sarasvathy's book, *Effectuation* (2008). She argues that entrepreneurs do not use managerial logic because they operate in unknowable environments. Rather than trying to predict the future, entrepreneurs work to control the future by co-creating it with self-selected stakeholders. They recognize the futility of predicting the future when the future is unknowable. For entrepreneurs, clarity comes through action and not analysis. They do not *discover* opportunities; they *create* opportunities. Saras calls this "effectual logic."

The notion that entrepreneurs get clarity through action turns the logic of strategy on its head. In Saras' world, execution would come before strategy. This is because an explicit recognition that the world is unknowable and that opportunities are made, not found, means that strategy does not follow vision; rather, action follows a vision that serves as a "soft goal" that will be shaped and molded into a strategy. Saras uses the analogy of cooking to make her point. One way to cook is to start with a clear goal of what one wants to eat. With a clear goal, execution means finding the right recipe for what we want to eat and then following it by making sure all the ingredients are at hand. The other way to cook is to start with a soft goal, "taking care of being hungry." There is, however, no particular dish or recipe in mind. We simply throw open the refrigerator and the spice cabinet, see what is at hand, and use these ingredients to throw together a dish. Note that both methods ultimately lead to the larger goal of nourishing one's body. The former begins with a goal and arrays the means to achieve the goal; the latter begins with the means at hand to achieve, and in the process shape and mold a larger vision.

The question of whether to use means-ends or ends-means logic to study strategy is a significant question for scholars of strategy and strategy execution. It raises the question of whether an organization that strictly follows the traditional causal logic for formulating and executing strategy is missing out on the possibility of innovation and creativity that is embedded in the logic of co-creation. With the massive failure of many of our major financial institutions, the validity of our traditional business models has acquired an added sense of urgency. It would be great for scholars of strategy to consider whether organizations that have successful strategy execution mechanisms of the type which Kaplan and Norton present in their book are also characterized by higher levels of creativity and innovation. This is an empirical question, and I do not have an answer, but I am convinced that it is a question worth studying. To the extent that this book brings the issue of organizational performance, creativity, opportunity creation, and its relationship to internal organizational structures and processes to the fore, it has accomplished an important purpose. For this we can thank Kaplan and Norton even if we do not find their entire message persuasive.

REFERENCES

Sarasvathy, S. D. 2008. *Effectuation: Elements of Entrepreneurial Expertise.* Cheltenham, U.K.: Edward Elgar Publishing.

SHAHID ANSARI
Professor of Behavioral and Managerial Accounting
Babson College

STEPHEN ZEFF (Editor), *Principles before Standards: The ICAEW's 'N Series' of Recommendations on Accounting Principles 1942–1969* (London, U.K.: ICAEW Financial Reporting Faculty, 2009, pp. 189).* No ISBN. Available in hard copy. For an electronic copy, go to www.icaew.com and search for Principles before standards.

The Institute of Chartered Accountants in England and Wales (ICAEW) was second only to the American Institute of Accountants in introducing a series of statements providing authoritative guidance on best accounting practice. Studying to become a member of the ICAEW in the late 1970s, I had to familiarize myself not only with the relatively new Statements of Standard Accounting Practice (SSAPs) being issued by the Accounting Standards Committee, but also with the "N Series," so-called after the section of the ICAEW's *Members' Handbook* in which the Recommendations on Accounting Principles were reproduced. The Recommendations were precisely as their name implied, guidance that the ICAEW, as it stated in preambles to various Recommendations, "hoped will be helpful to members when advising as to best practice." They were not mandatory, and hence preparers and auditors of financial statements could not be penalized if any Recommendations were not followed. Despite this, they provided an authoritative basis for financial reporting in the U.K., were used by the writers of textbooks as the basis for teaching accounting, and influenced both the SSAPs, which began to appear in 1971, and the requirements of U.K. company law relating to corporate financial statements.

The Recommendations form a central element of the history of standard setting for financial reporting, but because they are now superseded, it has been difficult for modern researchers and regulators to compare their early messages with more recent pronouncements. This collection has been published in the belief that the Recommendations are of interest in their own right as evidence of how professional accountants thought about financial reporting issues at a time when formal regulation was not a significant constraint on accounting policy choice. Readers are able to study not only the substantive recommendations on specific accounting practices, but also the quality of reasoning applied in developing the Recommendations. Over a period of some three decades, from 1942 to 1969, 29 Recommendations were published. The early Recommendations dealt with accounting for taxation, a matter made particularly complicated by almost confiscatory levels of business tax in the U.K. during World War II. To make matters worse, businesses paid income tax not on the profits of their current financial period, but rather on the profits of the period that finished in the preceding tax year (which for historical reasons ended on April 5). However, against the background of impending reform of U.K. company law in the mid-1940s, the Recommendations began to address more general issues, such as the presentation of the financial statements, reserves and provisions, accounting for business combinations, depreciation, and accounting for stock-in-trade (inventories). Many of the early Recommendations were incorporated into the Companies Act 1948.

The Recommendations went beyond annual financial statements to deal with prospectuses, estates and trusts, and retirement benefits. Several Recommendations were superseded by new pronouncements. For example, there were three substantially revised Recommendations on accounting for tax, stimulated by major reforms to the U.K. tax system. The Recommendations also touched on fundamental issues, with N12 (1949) and N15 (1952) both dealing with price change accounting. They recommended the continued use of historical cost, with any revaluations to take account of increased replacement costs being accounted for as reserves (equity) rather than as provisions (liabilities). Recommendations would respond to external events, with the ICAEW publishing a Recommendation on foreign currency translation (N25) in February 1968, just three months after the devaluation of Sterling in November 1967.

Stephen Zeff, who edited the collection, has supplemented the Recommendations themselves by providing a general introduction and a lengthy extract from his book, *Forging Accounting Principles in Five Countries* (1972). This sets out the complex and secretive process by which the ICAEW developed and approved the Recommendations. An important element of this was the involvement of ICAEW members who were not in public practice.

*The arrangements for this review were made by Christopher W. Nobes.

The proportion of members who were neither principals of nor employed by accountancy firms had been growing in the 1930s, and by the early 1940s over 20 percent of ICAEW members worked in business rather than practice. The Taxation and Financial Relations Committee, responsible for developing the Recommendations, was the first ICAEW committee to include members outside practice, and provided an impetus to greater involvement of such members in the management of the ICAEW, including eventually membership of Council. Zeff has also included an article, originally published in the weekly magazine *The Accountant* in 1953, in which Professor W. T. Baxter of the London School of Economics questioned whether professional accountancy bodies such as the ICAEW should make official pronouncements on "accounting theory" (Baxter objected to the support for historical cost given in N12 and N15, even though the documents themselves provided arguments for alternative accounting systems). Baxter was not a member of ICAEW, but rather a Scottish chartered accountant, and his own institute rejected the publication of similar recommendations. Baxter's opinion was that technical guidance could be useful, but that it should not replace professional judgment. He feared that the existence of formal statements of "approved" accounting theory would deter accounting students from exploring alternatives, and would ultimately lead to stagnation, as new ideas would have difficulty in challenging the "conventional wisdom."

Zeff also reproduces a hitherto unpublished interview with Michael Renshall, who had worked for the ICAEW as an administrator involved in the development of Recommendations since 1960, became the ICAEW's first Technical Director in 1969, and was active in the transition to SSAPs. Renshall provides insights into why the ICAEW considered it necessary to move beyond nonmandatory Recommendations, and notes the impact of financial scandals in the late 1960s. In his insider's account, Renshall celebrates the influence of a small number of key individuals (Ronald Leach of Peat, Marwick, Henry Benson of Coopers, and Professor Edward Stamp of Edinburgh University), suggesting that the creation of the Accounting Standards Steering Committee in 1969 owed a great deal to the personal charm and persuasiveness of Leach.

The re-publication of the Recommendations not only makes these important documents more accessible to students and researchers of accounting regulation but also provides examples of different styles of presenting guidance on financial reporting. The Recommendations range from a few paragraphs, presenting conclusions with little discussion, to lengthy explanations revealing the thought processes of their authors. Some of the Recommendations (for example, the two documents on accounting in conditions of changing prices) could be utilized in developing students' powers of reflection and critique, as they are at the same time concise and contentious. The supplementary items provide different angles on the standard-setting process. Does the emphasis on "due process," with standards being developed in full sight of the public, necessarily lead to higher quality standards than the secret processes of the 1950s and 1960s? Do standard setters have any legitimate authority to make pronouncements on accounting theory, or is their role at best one of applying principles developed by others to give guidance on technical accounting problems? Although the publication of *Principles before Standards* gives the ICAEW the opportunity of celebrating its contribution to the development of accounting standards, it also provides a valuable resource that will ensure that the "N Series" is appreciated by a wider readership.

REFERENCES

Baxter, W. T. 1953. Recommendations on accounting theory. *The Accountant* 129 (4112): 405–410.

Zeff, S. A. 1972. *Forging Accounting Principles in Five Countries: A History and an Analysis of Trends*. Champaign, IL: Stipes Publishing Co.

CHRISTOPHER NAPIER
Professor of Accounting
Royal Holloway, University of London

THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1479–1479

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1479

Errata

In the March 2010 issue of *The Accounting Review* article titled "Reporting Bias with an Audit Committee" by Judson Caskey, Venky Nagar, and Paolo Petacchi, the authors published an incomplete proof of Proposition 1 in the Appendix on page 469.[1] The proof of Proposition 1 of Caskey et al. (2010) was incomplete in its demonstration of the uniqueness of the linear equilibrium. The authors regret the oversight.

We first restate the equations that identify the equilibrium, which are listed as Equations (A8) and (A9) in the Appendix:

$$g_d \equiv p_f^3 + c_d^2\tau_y \text{var}(\text{E}[v|s_d,r_m])p_f - c_d^2\tau_y \text{var}(\text{E}[v|s_d,r_m]) = 0 \tag{1}$$

$$g_m \equiv r_x^3 + \tau_x \text{var}(\text{E}[v|s_m]|s_d)r_x - \frac{p_f}{c_m}\tau_x \text{cov}(v,\text{E}[v|s_m]|s_d) = 0. \tag{2}$$

The Appendix shows that a unique $p_f \in (0\,,\,1)$ solves (1) given r_x. When $r_x > 0$ the p_f that solves (1) is decreasing in r_x:

$$\frac{\text{d}p_f}{\text{d}r_x} = -\frac{\partial g_d/\partial r_x}{\partial g_d/\partial p_f} = -\frac{2c_d^2\dfrac{\tau_y}{\tau_x}\dfrac{\text{cov}(v,\text{E}[v|s_m]|s_d)^2}{(\text{var}(\text{E}[v|s_m]|s_d) + r_x^2\tau_x^{-1})^2}(1-p_f)r_x}{3p_f^2 + c_d^2\tau_y \text{var}(\text{E}[v|s_d,r_m])} < 0.$$

Our assumption that $\text{corr}(e_m\,,\,e_d) < \sqrt{\tau_m\,/\,\tau_d}$ implies that $\text{cov}(v\,,\,\text{E}[v \mid s_m] \mid s_d) > 0$. As a result, the unique p_f that solves (2) given r_x is increasing in r_x, crosses zero when $r_x = 0$, and approaches infinity as r_x approaches infinity. Thus, the p_f that solves (2) crosses the p_f that solves (1) at a unique point, giving the solution to Equations (1) and (2).

[1] Caskey, J., V. Nagar, and P. Petracchi. 2010. Reporting bias with an audit committee. *The Accounting Review* 85 (2): 447–481.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature. Topical areas of interest to the journal include accounting information systems, auditing and assurance services, financial accounting, management accounting, taxation, and all other areas of accounting, broadly defined. The journal is also open to all rigorous research methods.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and editors are responsible for providing constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be formatted in 12-point font on 8 1/2 x 11″ paper and should be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins should be at least one inch from top, bottom, and sides.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we".
5. A cover page should show the title of the paper, all authors' names, titles and affiliations, email addresses, and any acknowledgments.
6. The American Accounting Association encourages use of gender-neutral language in its publications.

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a cross-sectional equation; re-form*. See *Webster's* for correct usage.

Keywords: The abstract must be followed by at least three keywords to assist in indexing the paper and identifying qualified reviewers.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words (150 maximum) should be presented on a separate page immediately preceeding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. When the reference list contains two or more works by the same author (as the only author or first of two or more authors) published in the same year, the suffix a, b, etc., is appended to the date in the within-text citations and in the "References" section. For example, (Johansson 2004a, 2004b, 2004c; Baiman and Rajan 2002a, 2002b; Dhaliwal et al. 2005a, 2005b).
4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works should use acronyms or short titles where practicable. For example: (NCFFR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in *alphabetical order* according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Use authors' initials instead of proper names.
3. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
4. Date of publication follows the author's (authors') name(s).
5. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

———, and ———. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.

Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: http://hsgac.senate.gov/files/111803berry.pdf.

Cohen, D., A. Dey, and T. Lys. 2005. The Sarbanes Oxley Act of 2002: Implications for compensation structure and risk-taking incentives of CEOs. Working paper, New York University, University of Chicago, and Northwestern University.

Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition.* Ann Arbor, MI: University of Michigan.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

———, and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.

Dhaliwal, D., M. Erickson, and O. Li. 2005a. Shareholder income taxes and the relation between earnings and returns. *Contemporary Accounting Research* 22: 587–616.

———, L. Krull, O. Li, and W. Moser. 2005b. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43: 675–708.

Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.

Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109.* FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.

Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes.* Boston, MA: Houghton Mifflin.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.

Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics* (Winter) 30: 555–574.

National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: NCFFR.

Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.

———. 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy*, edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.

Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.

Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports.* Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be inserted using the "footnote" or "endnote" feature of the word processing software which will automatically number the footnotes throughout the manuscript with superscript Arabic numerals.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state upon submission that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. Manuscripts should be submitted via email as Microsoft Word or PDF file to the Senior Editor, Steven Kachelmeier, at email address tar@mccombs.utexas.edu. Please submit separate files for (1) the manuscript's title page with identifying information (not forwarded to reviewers), (2) the manuscript with title page and all other identifying information removed, and (3) any necessary supplemental files, such as experimental instructions and/or response memoranda on invited revisions. The body or subject line of the email should also indicate the submission fee receipt code from AAA (see step 4 below).
4. A nonrefundable submission fee is required unless explicitly waived by the editor for invited revisions of previous submissions. The submission fee in U.S. funds is $200.00 for members and $400.00 for nonmembers of the AAA payable by credit card (VISA or MasterCard only). The payment form is available online at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If you are unable to pay by credit card or have any questions, please contact the AAA Member Services Team at (941) 921-7747 or info@aaahq.org.
5. Revisions must be submitted within 12 months from the decision letter inviting a revision.

COMMENTS

The Accounting Review welcomes submissions of comments on previous *TAR* articles. Comments on articles previously published in *The Accounting Review* will generally be reviewed by two reviewers, usually including an author of the original article (to assist the editor in evaluating whether the submitted comment represents the prior article accurately) and an additional reviewer who is independent of the original article. If a comment is accepted for publication, the original author will generally be invited to reply. All other editorial requirements, as enumerated above, apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The AAA's Executive Committee policy (originally adopted in 1989, and amended in 2009) is that the objective of the Association-wide journals (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. To fulfill this objective, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles.

THE ACCOUNTING REVIEW

VOLUME 85, No. 5 SEPTEMBER 2010

ABHIJIT BARUA, STEVE LIN, and ANDREW M. SBARAGLIA
Earnings Management Using Discontinued Operations

BRIAN CADMAN, SANDY KLASA, and STEVE MATSUNAGA
Determinants of CEO Pay: A Comparison of ExecuComp and Non-ExecuComp Firms

PETER O. CHRISTENSEN, FLORIN ŞABAC, and JIE (JOYCE) TIAN
Ranking Performance Measures in Multi-Task Agencies

WILLIAM CREADY, THOMAS J. LOPEZ, and CRAIG A. SISNEROS
The Persistence and Market Valuation of Recurring Nonrecurring Items

MEI FENG and SARAH McVAY
Analysts' Incentives to Overweight Management Guidance When Revising Their Short-Term Earnings Forecasts

RICHARD C. HATFIELD, RICHARD W. HOUSTON, CHAD M. STEFANIAK, and SPENCER USREY
The Effect of Magnitude of Audit Difference and Prior Client Concessions on Negotiations of Proposed Adjustments

STEVEN J. KACHELMEIER and MICHAEL G. WILLIAMSON
Attracting Creativity: The Initial and Aggregate Effects of Contract Selection on Creativity-Weighted Productivity

PETRO LISOWSKY
Seeking Shelter: Empirically Modeling Tax Shelters Using Financial Statement Information

LILLIAN F. MILLS, LESLIE A. ROBINSON, and RICHARD C. SANSING
FIN 48 and Tax Compliance

TERENCE BU-PEOW NG and PREMILA GOWRI SHANKAR
Effects of Technical Department's Advice, Quality Assessment Standards, and Client Justifications on Auditors' Propensity to Accept Client-Preferred Accounting Methods

MARK E. PEECHER, M. DAVID PIERCEY, JAY S. RICH, and RICHARD M. TUBBS
The Effects of a Supervisor's Active Intervention in Subordinates' Judgments, Directional Goals, and Perceived Technical Knowledge Advantage on Audit Team Judgments

JACOB M. ROSE, CAROLYN STRAND NORMAN, and ANNA M. ROSE
Perceptions of Investment Risk Associated with Material Control Weakness Pervasiveness and Disclosure Detail

Marygrove College Library
8425 West McNichols Road
Detroit, MI 48221

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2010–2011 Executive Committee of the Association are as follows:

President	Kevin D. Stocks, Brigham Young University
President-Elect	Gregory B. Waymire, Emory University
Past President	Nancy A. Bagranoff, Old Dominion University
Vice President	Ira Solomon, University of Illinois at Urbana–Champaign
Vice President–Education	Belverd E. Needles, Jr., DePaul University
Vice President–Education-Elect	Alan Reinstein, Wayne State University
Vice President–Finance	Stacy E. Kovar, Kansas State University
Vice President–International	Kazuo Hiramatsu, Kwansei Gakuin University
Vice President–Professional Relations	Robert H. Colson, Grant Thornton, LLP
Vice President–Publications	Jean C. Bedard, Bentley University
Vice President–Publications-Elect	Christine A. Botosan, The University of Utah
Vice President–Research	John Christensen, University of Southern Denmark
Vice President–Sections and Regions	Gail Hoover King, Rockhurst University

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues include *Accounting Education News*, and online selection of one, two, or all three of the following Association-wide journals: *The Accounting Review*, *Accounting Horizons*, and *Issues in Accounting Education*. (This does not include Section journals.) Select from:

The Accounting Review only	U.S. $185.00	with Print $220.00
Accounting Horizons OR *Issues in Accounting Education* only	U.S. $185.00	with Print $200.00
The Accounting Review AND *Accounting Horizons* OR *Issues in Accounting Education*	U.S. $195.00	with Print $245.00
Accounting Horizons AND *Issues in Accounting Education*	U.S. $195.00	with Print $225.00
All three Association-wide journals	U.S. $205.00	with Print $270.00

Associate membership:
Full-time students are eligible for associate membership. The print option is not available for Associate members. Select from:

Any one Association-wide journal	U.S. $25.00
Any two Association-wide journals	U.S. $35.00
All three Association-wide journals	U.S. $45.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Auditing; Diversity; Financial Accounting and Reporting; Forensic and Investigative Accounting; Gender Issues and Worklife Balance; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Strategic and Emerging Technologies; Teaching, Learning and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to tar@mccombs.utexas.edu. If it is necessary to send any hard copy correspondence, the mailing address is Mary Capps, Editorial Assistant, *The Accounting Review*, The University of Texas at Austin; Department of Accounting, CBA 4M202; 1 University Station B6400; Austin, TX 78712-0211 • Telephone: (512) 471-5318 • Fax: (512) 471-3904.

The submission fee of $200.00 for members of the AAA or $400.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Info@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review is indexed in the Social Science Citation Index which is included in Web of Science.

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rates of $375 for a print copy or at $425 for a print and electronic copy. The member subscription rate for a hard copy is $35. Electronic access is included with membership dues. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review*, 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2010 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Steven J. Kachelmeier

Editors

Shannon W. Anderson
J. Harry Evans III
James E. Hunton
Kathryn Kadous
Sanjay G. Kallapur
Ranjani Krishnan
Laureen A. Maines
D. Paul Newman
Thomas C. Omer
Wayne B. Thomas
Mark A. Trombley
Beverly R. Walther
Paul A. Zarowin

VOL. 85 SEPTEMBER 2010 No. 5

ABHIJIT BARUA, STEVE LIN, and ANDREW M. SBARAGLIA
Earnings Management Using Discontinued Operations 1485
BRIAN CADMAN, SANDY KLASA, and STEVE MATSUNAGA
Determinants of CEO Pay: A Comparison of ExecuComp and Non-ExecuComp Firms 1511
PETER O. CHRISTENSEN, FLORIN ŞABAC, and JIE (JOYCE) TIAN
Ranking Performance Measures in Multi-Task Agencies 1545
WILLIAM CREADY, THOMAS J. LOPEZ, and CRAIG A. SISNEROS
The Persistence and Market Valuation of Recurring Nonrecurring Items 1577
MEI FENG and SARAH McVAY
Analysts' Incentives to Overweight Management Guidance When Revising Their Short-Term Earnings Forecasts 1617
RICHARD C. HATFIELD, RICHARD W. HOUSTON, CHAD M. STEFANIAK, and SPENCER USREY
The Effect of Magnitude of Audit Difference and Prior Client Concessions on Negotiations of Proposed Adjustments 1647
STEVEN J. KACHELMEIER and MICHAEL G. WILLIAMSON
Attracting Creativity: The Initial and Aggregate Effects of Contract Selection on Creativity-Weighted Productivity 1669
PETRO LISOWSKY
Seeking Shelter: Empirically Modeling Tax Shelters Using Financial Statement Information 1693
LILLIAN F. MILLS, LESLIE A. ROBINSON, and RICHARD C. SANSING
FIN 48 and Tax Compliance 1721
TERENCE BU-PEOW NG and PREMILA GOWRI SHANKAR
Effects of Technical Department's Advice, Quality Assessment Standards, and Client Justifications on Auditors' Propensity to Accept Client-Preferred Accounting Methods 1743
MARK E. PEECHER, M. DAVID PIERCEY, JAY S. RICH, and RICHARD M. TUBBS
The Effects of a Supervisor's Active Intervention in Subordinates' Judgments, Directional Goals, and Perceived Technical Knowledge Advantage on Audit Team Judgments 1763
JACOB M. ROSE, CAROLYN STRAND NORMAN, and ANNA M. ROSE
Perceptions of Investment Risk Associated with Material Control Weakness Pervasiveness and Disclosure Detail 1787

BOOK REVIEWS, Stephen A. Zeff, Editor

Peter O. Christensen and Gerald A. Feltham
Equity Valuation ..... XIAO-JUN ZHANG **1809**

Paul M. Clikeman
Called to Account: Fourteen Financial Frauds that Shaped the American Accounting Profession ..... W. MORLEY LEMON **1811**

E. Richard Criscione
Abraham J. (Abe) Briloff: A Biography ..... GEORGE FOSTER **1814**

Werner De Bondt (editor)
Financial Accounting and Investment Management ..... STEVEN J. MONAHAN **1816**

Ken McPhail and Diane Walters
Accounting and Business Ethics: An Introduction ..... JAMES C. GAA **1817**

Othmar W. Winkler
Interpreting Economic and Social Data: A Foundation of Descriptive Statistics ..... THOMAS R. DYCKMAN **1820**

Editorial Policy and Style Information **1823**

Editorial Staff of *The Accounting Review*
2010–2011

SENIOR EDITOR
Steven J. Kachelmeier, The University of Texas at Austin

EDITORS
Shannon W. Anderson, University of California, Davis
J. Harry Evans III, University of Pittsburgh
James E. Hunton, Bentley University
Kathryn Kadous, Emory University
Sanjay G. Kallapur, Indian School of Business
Ranjani Krishnan, Michigan State University
Laureen A. Maines, Indiana University
D. Paul Newman, The University of Texas at Austin
Thomas C. Omer, Texas A&M University
Wayne B. Thomas, University of Oklahoma
Mark A. Trombley, The University of Arizona
Beverly R. Walther, Northwestern University
Paul A. Zarowin, New York University

EDITORIAL ASSISTANT
Mary C. Capps, The University of Texas at Austin

RESEARCH ASSISTANT
Tracie Majors, The University of Texas at Austin

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland

COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Lisa Habblitz
Lindsay Wharton
Kathy Casper
Judy Cothern
Sonja Langley
Jan Kovarik (subcontracted)
Copyediting Team (subcontracted)

OTHER AAA STAFF
Dierdre Harris
David Boynton
Alex Metz
Pat Stein
Debora Gardner
Peggy Turczyn
Barbara Gutierrez
Barbara Brady
Arlene Oliver
Kathleen Messenger
Roland LaTulip
Diane Leger
Man Fong
Jodie M. Morrill
Michelle Bowles

Editorial Advisory and Review Board

Anwer S. Ahmed, Texas A&M University
Ashiq Ali, The University of Texas at Dallas
Anil Arya, The Ohio State University
Benjamin C. Ayers, University of Georgia
William R. Baber, Georgetown University
Stephen P. Baginski, University of Georgia
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Linda Smith Bamber, University of Georgia
Orie E. Barron, The Pennsylvania State University
Anne L. Beatty, The Ohio State University
Jean C. Bedard, Bentley University
Robert J. Bloomfield, Cornell University
Jennifer L. Blouin, University of Pennsylvania
Mark T. Bradshaw, Boston College
Lawrence D. Brown, Georgia State University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
Joseph V. Carcello, The University of Tennessee
Mary Ellen Carter, Boston College
Paul K. Chaney, Vanderbilt University
Qiang Cheng, University of Wisconsin–Madison
Michael B. Clement, The University of Texas at Austin
William M. Cready, The University of Texas at Dallas
Antonio Davila, IESE Business School–University of Navarra
Patricia M. Dechow, University of California, Berkeley
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Sunil Dutta, University of California, Berkeley
Leslie G. Eldenburg, The University of Arizona
Patricia M. Fairfield, Georgetown University
David B. Farber, University of Missouri–Columbia
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jére R. Francis, University of Missouri–Columbia
James R. Frederickson, Melbourne Business School
Frank B. Gigler, University of Minnesota
Cristi A. Gleason, The University of Iowa
Johnathan C. Glover, Carnegie Mellon University
Steven M. Glover, Brigham Young University
David A. Guenther, University of Oregon
Luzi Hail, University of Pennsylvania
Jeffrey Hales, Georgia Institute of Technology
John R. M. Hand, The University of North Carolina at Chapel Hill
R. Lynn Hannan, Georgia State University
Gilles Hilary, INSEAD
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
Chris E. Hogan, Michigan State University
Ole-Kristian Hope, University of Toronto
Richard W. Houston, The University of Alabama
Raffi J. Indjejikian, University of Michigan
Christopher D. Ittner, University of Pennsylvania
Debra C. Jeter, Vanderbilt University
Karla M. Johnstone, University of Wisconsin–Madison
Bin Ke, Nanyang Technological University
S. Jane Kennedy, University of Washington
Inder K. Khurana, University of Missouri–Columbia
Michael D. Kimbrough, University of Maryland

Ronald R. King, Washington University in St. Louis
April Klein, New York University
W. Robert Knechel, University of Florida
Lisa L. Koonce, The University of Texas at Austin
Susan D. Krische, American University
Jayanthi Krishnan, Temple University
William N. Lanen, University of Michigan
David F. Larcker, Stanford University
Reuven Lehavy, University of Michigan
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Joan L. Luft, Michigan State University
Anne M. Magro, George Mason University
Carol Marquardt, Baruch College–CUNY
Dawn A. Matsumoto, University of Washington
Edward Maydew, The University of North Carolina at Chapel Hill
Brian W. Mayhew, University of Wisconsin–Madison
Sarah E. McVay, University of Utah
Kenneth A. Merchant, University of Southern California
Gregory S. Miller, University of Michigan
Lillian F. Mills, The University of Texas at Austin
Brian Mittendorf, The Ohio State University
Partha S. Mohanram, University of Toronto
Steven J. Monahan, INSEAD
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Arkansas
Venky Nagar, University of Michigan
Terry L. Neal, The University of Tennessee
Karen K. Nelson, Rice University
Edward F. O'Donnell, Southern Illinois University Carbondale
Per M. Olsson, Duke University
Chul W. Park, The University of Hong Kong
Mark E. Peecher, University of Illinois at Urbana–Champaign
Sean A. Peffer, University of Kentucky
Kathy Ruby Petroni, Michigan State University
John D. Phillips, University of Connecticut
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, Stanford University
Marlene A. Plumlee, University of Utah
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shiva Rajgopal, Emory University
Scott Richardson, BlackRock Inc.
John R. Robinson, The University of Texas at Austin
Darren T. Roulstone, The Ohio State University
Katherine Schipper, Duke University
Karen L. Sedatole, Michigan State University
Catherine M. Shakespeare, University of Michigan
Lakshmanan Shivakumar, London Business School
K. (Shiva) Sivaramakrishnan, University of Houston
Hollis A. Skaife, University of Wisconsin–Madison
Richard G. Sloan, University of California, Berkeley
J. Reed Smith, Indiana University at Indianapolis
Mark T. Soliman, University of Washington
Geoffrey B. Sprinkle, Indiana University
Mary Harris Stanford, Texas Christian University
Phillip C. Stocken, Dartmouth College
K. R. Subramanyam, University of Southern California
Hun-Tong Tan, Nanyang Technological University
Kristy L. Towry, Emory University
Senyo Tse, Texas A&M University
Jennifer Wu Tucker, University of Florida
Irem Tuna, London Business School
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Michael S. Wilkins, Texas A&M University
Richard H. Willis, Vanderbilt University
Arnold M. Wright, Northeastern University
Teri Lombardi Yohn, Indiana University
Xiao-Jun Zhang, University of California, Berkeley
Mark F. Zimbelman, Brigham Young University

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1959–1962	Robert K. Mautz
1962–1965	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William Kinney
1990–1994	A. Rashad Abdel-khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber
2002–2005	Terry Shevlin
2005–2008	Dan S. Dhaliwal

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1485–1509

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1485

Earnings Management Using Discontinued Operations

Abhijit Barua
Steve Lin
Andrew M. Sbaraglia
Florida International University

ABSTRACT: This study investigates whether managers use classification shifting to manage earnings when reporting discontinued operations. Using a methodology similar to McVay (2006), we find evidence consistent with the hypothesis that firms shift operating expenses to income-decreasing discontinued operations to increase core earnings. Our findings also indicate that managers use classification shifting to meet or beat analysts' forecasts. Finally, we find that, since the introduction of SFAS No. 144, the reporting frequency of discontinued operations has increased; however, the magnitude of classification shifting has decreased. We provide potential explanations for this finding.

Keywords: *earnings management; classification shifting; earnings components; discontinued operations.*

Data Availability: *Data are available from public sources identified in this study.*

I. INTRODUCTION

Classification shifting involves reporting revenues, expenses, gains, and losses on different lines on the income statement other than those on which they should properly appear under GAAP. In particular, this study focuses on whether managers achieve classification shifting by allocating operating expenses to discontinued operations. Managers could engage in this form of earnings management if they perceive that investors value recurring earnings more than nonrecurring earnings. It is difficult for investors to detect this type of earnings management because discontinued operations are not usually disclosed in detail in the financial statements. Unlike other forms of earnings management, such as accrual management or real activity manipulation, classification shifting does not change net income and may be a less costly form of earnings management.

We appreciate the constructive comments and suggestions from Steven Kachelmeier, Mark Kohlbeck, Sarah McVay (discussant), Paul Zarowin, two anonymous reviewers, and the seminar participants at the 2009 AAA Annual Meeting, 2008 Southeast Regional AAA Conference, Florida International University, Florida Atlantic University, Tainan University of Technology (Taiwan), Gakushuin University (Japan), and Waseda University (Japan). All remaining errors are our own.

Editor's note: Accepted by Paul Zarowin.

Submitted: August 2008
Accepted: December 2009
Published Online: August 2010

We believe this is an interesting question for the following reasons. First, previous studies have provided evidence of classification shifting using extraordinary items (Ronen and Sadan 1975; Barnea et al. 1976) and special items (McVay 2006; Fan et al. 2010), while discontinued operations have not been studied. It could be more beneficial for managers to use discontinued operations for classification shifting than special items, because shifting operating expenses to discontinued operations also increases operating income, income from continuing operations, and core earnings.[1] Increasing all three subtotals could lead to higher stock valuation than increasing core earnings alone.[2] Second, there is no detailed disclosure of discontinued operations in the financial statements. This lack of detail creates information asymmetry between managers and investors, as investors do not know the nature of the expenses allocated to discontinued operations. Third, the adoption of SFAS No. 144, *Accounting for the Impairment or Disposal of Long-Lived Assets* (FASB 2001a), has increased the reporting frequency of discontinued operations by reducing the recognition threshold for asset disposals to be classified as discontinued operations. Consequently, the FASB may have inadvertently allowed managers more discretion to manage earnings, potentially reducing the quality of reported earnings. Finally, in recent years, special items have been scrutinized by researchers, regulators, and auditors, while discontinued operations have been largely ignored. This increased scrutiny may have increased the costs and reduced the benefits of classification shifting using special items. It could now be more advantageous for managers to use discontinued operations, especially since the recognition threshold has been reduced since SFAS No. 144. To our knowledge, ours is the first study to investigate the implication of increasing recognition of discontinued operations for earnings management after the introduction of SFAS No. 144. This study benefits investors, analysts, auditors, and regulators by alerting them to the potential earnings management using classification shifting when a firm reports discontinued operations.

Using a research methodology similar to McVay (2006), we decompose core earnings into its expected and unexpected components. We find a positive association between income-decreasing discontinued operations and unexpected core earnings in the year a firm reports discontinued operations, a result consistent with managers shifting operating expenses to discontinued operations.[3] We also find a negative association between discontinued operations and the change in unexpected core earnings in the year after a firm reports discontinued operations, a result consistent with operating expenses returning to core earnings. Together, these results provide evidence that managers shift core expenses to discontinued operations to inflate core earnings. We also investigate the motivations for this form of classification shifting, finding evidence consistent with managers using classification shifting to meet or beat analysts' forecasts.

Our contribution to the literature is threefold. First, we extend the classification shifting literature by providing evidence that discontinued operations are used to manage earnings. Second, we extend the benchmark-beating literature by providing evidence that managers engage in classification shifting using discontinued operations to meet or beat analyst forecasts. Finally, we

---

[1] For ease of exposition, we characterize classification shifting as the shifting of core expenses to discontinued operations. It is also possible that managers shift revenues from discontinued operations to core earnings; both of these shifting mechanisms are captured by our examination of classification shifting.

[2] While using extraordinary items for classification shifting yields the same benefits as using discontinued operations, they are difficult to recognize. Over time, the FASB has increasingly tightened the standards for reporting extraordinary items. Now, even catastrophic events such as the attack on September 11, 2001 and Hurricane Katrina are not considered extraordinary events for financial reporting purposes (EITF 01-10, FASB 2001c). In 2003, only 2 percent of firms included in *Accounting Trends and Techniques,* 60th edition, reported extraordinary items.

[3] The classification shifting increases unexpected core earnings and reduces the value of the discontinued operations creating a negative relationship. We multiply the value of discontinued operations by negative one and predict a positive relationship.

document an increase in the reporting frequency of discontinued operations and a decrease in the magnitude of earnings management since the introduction of SFAS No. 144. There are two potential reasons for the last finding. First, SFAS No. 144 broadened the definition of discontinued operations by replacing the "business segment" requirement under APB No. 30 with the "component of an entity" concept. This allowed for more asset dispositions to be accounted for as discontinued operations, increasing the reporting frequency. In addition, the broadened definition also allowed smaller asset dispositions to qualify as discontinued operations, potentially reducing the amount of operating expenses that could be allocated to the disposition. Second, the Sarbanes-Oxley Act was enacted around the same time as SFAS No. 144, and could have led to a decrease in classification shifting due to increased regulation and oversight.[4]

Section II discusses discontinued operations and classification shifting. Section III reviews prior literature and develops testable hypotheses. Section IV describes the sample selection criteria, data collection method, and variable measurement. Section V reports the results, and Section VI concludes.

II. DISCONTINUED OPERATIONS AND CLASSIFICATION SHIFTING

When making firm-valuation decisions, investors place a higher value on items of income expected to be persistent in the future. Presumably to aid users' valuation decisions, GAAP generally requires that material nonrecurring items be separately disclosed in the financial statements. However, the separation of net income into recurring and nonrecurring components also gives managers the opportunity to mislead investors, by misclassifying income and expense items. For example, a manager wanting to increase firm valuation can misclassify recurring expenses as nonrecurring, misleading investors as to the persistence of the income increase. In the accounting literature, this type of earnings management is called classification shifting.

McVay (2006) discusses two reasons why classification shifting can be a relatively low-cost method for managing earnings. First, unlike accrual management or real activity manipulation, there is no "settling-up" in the future for past earnings management. If a manager decides to increase earnings using income-increasing accruals, then, at some point, these accruals must reverse. The reversal of these accruals reduces future reported earnings. If a manager decides to increase earnings by managing real activities, such as reducing research and development expenditures, then this likely leads to fewer income-producing projects and reduced earnings in the future. In contrast, classification shifting involves simply reporting recurring expenses in a nonrecurring classification on the income statement, having no implications for future earnings. Second, because classification shifting does not change net income, it is potentially subject to less scrutiny by auditors and regulators than forms of earnings management that change net income.

Discontinued operations represent the income and cash flows of a portion of a company that has been (or will be) discontinued from the continuing operations of the company. Under SFAS No. 144, the gain or loss for discontinued operations is comprised of three amounts. The first amount is the operating income or loss from the operations of the component being discontinued for the entire year in which the decision to discontinue is made. The second amount is the gain or loss from disposal, which is the net amount realized over the carrying value of net assets of the component. The third amount is an impairment loss on the "assets held for sale" if the component is not disposed in the same year as the decision to discontinue.[5] The results of discontinued

[4] Lobo and Zhou (2006) find that accounting conservatism has increased in the post-Sarbanes-Oxley period. In addition, Cohen et al. (2008) find that accrual-based earnings management has decreased following SOX, while real activity management has increased.

[5] The aggregate gain or loss from discontinued operations was computed differently under APB No. 30. The income or loss from operations of the segment being disposed was measured as the income or loss from the operation from the

operations are reported on the income statement as a separate item (net of tax effects) after income from continuing operations. The "below the line" reporting of disposals as discontinued operations is likely perceived to be desirable for investors, because it communicates to them the results of a firm's operations on a "with and without" basis. However, it could be detrimental to investors if managers use their reporting discretion to manage earnings when reporting discontinued operations. For example, managers could allocate normal operating expenses to discontinued operations in order to report increased income from continuing operations. Managers have discretion over the allocation of joint costs between continuing and discontinued operations, and this information is not normally available to external investors. It is doubtful that auditors would detect this manipulation during their analytical review because the company's financial ratios would be similar to what they were previously.[6]

The ability of asset disposals to be classified as discontinued operations has varied throughout time. Under APB Opinion No. 30 (APB 1973), only dispositions of business "segments" were able to qualify for reporting as discontinued operations. In general, business segments were defined as a major line of business or a customer class. The FASB believed that investors would benefit from expanded application of the accounting and disclosure rules of discontinued operations and issued SFAS No. 144, *Accounting for the Impairment or Disposal of Long-Lived Assets* (FASB 2001a). SFAS No. 144 reduced the threshold for recognition of discontinued operations treatment by introducing the "component of an entity" concept. A component of an entity is distinguishable from the rest of the entity because it has its own clearly defined operations and cash flows. A component of an entity can be a reportable segment, an operating segment (as defined by SFAS No. 131, FASB 1997), a reporting unit (as defined by SFAS No. 142, FASB 2001b), a subsidiary, or an asset group. This reduction in the threshold for discontinued operations has allowed for more asset disposals to be classified as discontinued operations. Consequently, the FASB may have inadvertently given managers more discretion to manage earnings, potentially reducing the quality of reported earnings.

In addition to the "component of an entity" requirement, SFAS No. 144 mandated two other requirements for an asset disposal to qualify for treatment as a discontinued operation. The first requirement is that the operations and cash flows of the component being disposed are completely separable from the seller's continuing operations. The second requirement is that the seller has no significant involvement with the component after the sale. In theory, the application of these requirements appears straightforward; however, SFAS No. 144 did not clearly define the meaning of significant involvement. In practice, many asset disposals involve contractual terms: seller financing, retained equity by the seller, royalties paid to the seller by the buyer, service agreements, etc. The inclusion of these terms in the asset-disposal contract could be interpreted as significant involvement by the seller. In order to provide better guidance on the meaning of the term "significant involvement," the FASB issued EITF 03-13 (FASB 2004) as a guide for determining whether an entity retains significant involvement or receives cash flows. In short, if a firm receives material direct cash flows from the discontinued operations or maintains significant con-

---

beginning of the year until the "measurement date." The measurement date was the date the plan to dispose of the segment was approved by the company. If the disposal took place in the same year as the measurement date, then the gain or loss on disposal was comprised of two items. The first was the gain or loss from operations from the measurement date to the end of the year and the second was the difference between the selling price of the segment and its book value. If the segment was not disposed in the same year, then the gain or loss on disposal also included an estimate of future losses.

[6] For example, assume a retailer is closing a number of individual stores and accounts for this as a discontinued operation. Further assume the retailer always allocated a percentage of advertising expenses to these stores and plans no reduction in advertising expense after the stores are closed. In this case, the advertising expense for the continuing operations of the retailer will be the same on a per-store basis as in the past.

tinuing involvement, then the disposal does not qualify as a discontinued operation. EITF 03-13 took effect in 2005 but could be adopted as early as 2004. It should be noted that, while the recognition criteria for discontinued operations are well defined in GAAP, the costs to be allocated to the discontinued operations are not.

Accounting treatment for discontinued operations is also a global accounting issue. Similar to APB No. 30, IFRS 5, *Non-Current Assets Held-for-Sale and Discontinued Operations* (IASB 2004) defines a discontinued operation as a separate major line of business or geographical area of operations. IFRS 5 also requires detailed disclosure of revenue, expenses, pre-tax profit or loss, and the related income tax expense, either in the notes or on the face of the income statement. Currently, both the FASB and the IASB are working toward a converged accounting definition and treatment for discontinued operations, with both boards issuing proposals amending SFAS No. 144 and IFRS 5, respectively.

In closing, material asset disposals are reported on the income statement as discontinued operations if they meet the requirements of SFAS No. 144 (or APB No. 30 for disposals prior to 2002) and likely reported as special items if they do not. Either income statement classification gives managers the opportunity to engage in classification shifting. Prior research provides evidence that special items are used for classification shifting. Since the recognition criteria for discontinued operations are better defined in the accounting standards than for special items, this may give managers less flexibility to manage earnings using classification shifting. For example, it may not be possible for managers to time the recognition of discontinued operations to coincide with a period during which earnings management is desired. Consequently, whether managers use discontinued operations for earnings management purposes is an empirical question that our research addresses

III. LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

Earnings management is one of the most studied areas in financial accounting research.[7] Previous studies document that earnings management can be carried out through accrual management (e.g., Dechow et al.1995; Payne and Robb 2000), real activity management (e.g., Dechow and Sloan 1991; Bushee 1998; Roychowdhury 2006),[8] or classification shifting (Ronen and Sadan 1975; Barnea et al. 1976; McVay 2006; Fan et al. 2010). However, increasing current earnings using the former two methods has the potential of reducing future earnings. Since classification shifting simply moves certain revenues, expenses, gains, and losses to different line items on the income statement, it does not actually change net income. As a result, classification shifting is likely to be less costly and less scrutinized by auditors and regulators (Nelson et al. 2002).[9] Previous studies have largely focused on earnings management using accruals and real activity management. Relatively few studies have examined earnings management through classification shifting.

---

[7] A review of the accrual-based earnings management literature can be found in Healy and Wahlen (1999), Dechow and Skinner (2000), and Kothari (2001). Xu et al. (2007) review the real activity-based earnings management literature.

[8] Discretionary expenses may include research and development (Dechow and Sloan 1991; Bushee 1998), and advertising, maintenance, and training expenses (Roychowdhury 2006).

[9] Nelson et al. (2002) investigate auditors' responses to earnings management attempts. Their survey evidence suggests that auditors are less likely to require audit adjustments under the following circumstances: first, when GAAP is specific and the firm structures the transaction to comply with GAAP; second, when GAAP is not specific and the firm engages in an unstructured transaction; third, when the transaction does not increase current income; fourth, when the client is large; and fifth, when the amount of the earnings management is immaterial. For a transaction to be categorized as discontinued operation, the firm needs to structure the transaction to comply with specific accounting standards, which would presumably reduce the chance of an audit adjustment. The calculation of the income from the discontinued operations is rather unstructured, which would reduce the probability of an audit adjustment. Because classification shifting does not increase net income, their results would again suggest the reduced probability of an audit adjustment. Finally, the large size of our sample firms would also reduce the likelihood of an audit adjustment.

Ronen and Sadan (1975) contend that the dimensions and objects of income smoothing are closely associated. They also reason that, if the smoothing object is net income, then classification shifting is irrelevant. However, if the smoothing object is any income subtotal other than "bottom-line" net income, then managers have an incentive to engage in classification shifting. This begs the question of which income subtotals interest investors. Lipe (1986) documents that investors understand the future earnings implications between the different earnings components reported on the income statement. Bradshaw and Sloan (2002) provide evidence of "street earnings" (i.e., modified-GAAP earnings with noncash and nonrecurring items excluded) replacing GAAP earnings as one of the primary determinants of stock price. Together, these studies provide evidence of investors' interest in recurring income subtotals rather than "bottom line" net income that includes nonrecurring items. If investors value recurring earnings higher than nonrecurring earnings, then managers have an incentive to misclassify operating expenses as nonrecurring expenses, to increase recurring income subtotals. Although one cannot observe which income statement subtotals investors use, we follow McVay (2006) in assuming that core earnings is the managed subtotal. However, using discontinued operations to increase core earnings also increases income from continuing operations and other "above-the-line" earnings subtotals.

One of the earliest studies in the area of classification shifting is Ronen and Sadan (1975). They find evidence consistent with managers using extraordinary items to smooth earnings before extraordinary items. Barnea et al. (1976) extend this line of research by providing evidence that managers also use extraordinary items to smooth operating income. Taken together, these studies provide evidence that managers are interested in managing multiple income statement subtotals, presumably because investors value various income statement subtotals differently. While the above studies provide evidence of earnings management using "below the line" income statement items, they do not examine the motivations behind classification shifting.

Kinney and Trezevant (1997) investigate whether managers use special items opportunistically to manage both earnings and investors' perceptions. They find that income-decreasing special items are more likely to be reported as line items on the income statement, presumably to call attention to the transitory nature of the earnings decrease caused by these items. They also find that income-increasing special items are more likely to be reported in the footnotes to the financial statements, presumably to shift attention away from the transitory nature of the earnings increase caused by these items. Their findings suggest that managers behave opportunistically when disclosing special items. However, Riedl and Srinivasan (2010) further examine managers' reporting behavior regarding special items. They find that special items separately disclosed on the income statement have a lower persistence than special items disclosed in the financial statement footnotes. Their results are consistent with managers using financial statement presentation to assist investors in evaluating the persistence of special items. This evidence offers an alternative explanation for the managerial opportunism detected by Kinney and Trezevant (1997).

McVay (2006) investigates whether managers engage in classification shifting, by reporting core expenses in income-decreasing special items. She uses an expectation model to separate core earnings, defined as operating income before depreciation and amortization, into expected and unexpected components. She finds special items are positively associated with contemporaneous unexpected core earnings and negatively associated with the unexpected change in future core earnings. These results are consistent with her hypothesis that managers shift operating expenses to special items. She also provides evidence suggesting that managers' motivation for classification shifting is to meet or beat analysts' forecasts. In addition, she finds a negative stock price reaction to the reversal of unexpected core earnings, indicating that investors do not understand (or are not able to fully undo) the earnings management. One drawback to the core earnings expectation model used by McVay (2006) is the use of contemporaneous accruals (including accruals

related to special items) as a control for firm performance. The inclusion of special item accruals in the expectation model creates a potential bias in favor of her hypotheses. She acknowledges the reliance on an imperfect model is a limitation of her study.

Fan et al. (2010) extends McVay's (2006) core earnings expectation models by controlling for performance using returns and lagged returns rather than contemporaneous accruals, thus eliminating any potential bias. They confirm McVay's (2006) findings that managers shift core expenses to income-decreasing special items. Using quarterly data, they find that classification shifting using income-decreasing special items is more prevalent in the fourth quarter than in other quarters. They also find that classification shifting is more prevalent when managers are constrained from using income-increasing accruals.

Classification Shifting Using Discontinued Operations

This study investigates whether discontinued operations are used for classification shifting. Our study is motivated by prior research suggesting that managers are opportunistic when reporting "below the line" income statement items. Results from discontinued operations are reported at a lower point on the income statement than special items, but are arguably subject to less management discretion. Thus, while managers have a greater incentive to shift expenses to this classification, it is an empirical question as to whether they have the opportunity to do so. In addition, shifting operating expenses to discontinued operations also increases core earnings, operating income, and income from continuing operations. These subtotals are widely believed to be good indicators of firm performance and to be of primary interest to investors. Discontinued operations could presumably attract less attention from investors, auditors, and regulators than special items.

Discontinued operations are reported on the income statement in an aggregated form, creating an information asymmetry between managers and investors. It is unclear to investors exactly what revenues, expenses, gains, and losses are allocated to the discontinued operations. Because of this, we predict a shifting of operating expenses to discontinued operations. This leads to our first hypothesis, stated in alternative form:

H1: Managers engage in classification shifting using discontinued operations to increase core earnings.

Classification Shifting to Meet or Beat Earnings Benchmarks

The literature provides evidence of firms managing earnings to meet or beat three earnings benchmarks: zero earnings, prior year's earnings, and analysts' forecasts. Prior research also shows that firms have several methods of managing earnings to meet these benchmarks. For example, prior research documents a discontinuity in the earnings distribution, with more firms reporting income just above these benchmarks than firms reporting income below these benchmarks (Hayn 1995; Burgstahler and Dichev 1997; Degeorge et al. 1999). Degeorge et al. (1999) reason this is due to certain investors relying on simple heuristics when evaluating firm performance. They further reason that because investors are risk-averse, they have a tendency to be more sensitive toward losses than toward gains. The literature has also documented market rewards for firms exceeding financial benchmarks. Barth et al. (1999) find firms that consistently exceed prior years' earnings have higher price-earnings multiples than firms that do not. Brown and Caylor (2005) find that, even after controlling for the earnings surprise, firms experience positive abnormal returns when reporting positive earnings, earnings increases, or exceeding analysts' forecasts. Together, these studies suggest that firms exceeding benchmarks have a lower cost of equity. Jiang (2008) documents that firms exceeding benchmarks also have a lower cost of debt, especially for firms exceeding the zero earnings benchmark. Prior literature has also provided evidence that

managers receive rewards for exceeding benchmarks. For example, Matsunaga and Park (2001) find evidence consistent with CEOs receiving a lower cash bonus when their firms report an earnings decline or miss analysts' forecasts.

We have no prediction as to which benchmark is the most important to firms reporting discontinued operations. Following Degeorge et al. (1999), Brown and Caylor (2005), and Jiang (2008), we conjecture that managers use classification shifting to meet or beat the following earnings benchmarks: zero earnings (avoiding a loss), prior year's earnings (avoiding an earnings decline), and exceeding the consensus analyst forecast (avoiding a negative earnings surprise). More formally, we predict the following:

H2a: Managers engage in classification shifting using discontinued operations to increase core earnings in order to avoid reporting negative core earnings.

H2b: Managers engage in classification shifting using discontinued operations to increase core earnings in order to meet or beat prior year's core earnings.

H2c: Managers engage in classification shifting using discontinued operations to increase core earnings in order to meet or beat analyst forecasts.

Classification shifting using discontinued operations involves two managerial decisions. The first is the decision to discontinue the operations, while the second is to shift operating expenses to this classification. We argue that managers have already made the decision to discontinue certain operations before deciding to engage in classification shifting. In other words, we believe the two decisions are independent of each other. We base our belief on the following reasoning. First, discontinued operations are an important real activity decision requiring board of director approval. Consequently, the manager needs compelling economic reasons to make a case to the board. Second, because discontinued operations are a real activity decision, the company will incur real expenses disposing of the operation, paying expenses such as severance pay, appraisal fees, legal fees, etc. from the cash flows of continuing operations. Third, there is likely to be a possibility of improvement in the operating results of the discontinued operation. If so, the manager will be trading future earnings to increase current period earnings. Finally, we believe that the decision to discontinue an operation is a lengthy process, requiring board approval, asset appraisals, identification of potential purchasers, etc.

In sum the decision to discontinue an operation is both an expensive and time-consuming decision. If made solely for the purpose of earnings management, then it would reduce the main advantages of classification shifting, namely being relatively simple and inexpensive. This study focuses on the financial reporting decision rather than on the decision to discontinue.

IV. DATA, SAMPLE, AND VARIABLE MEASUREMENT

The sample used in this study is comprised of data from the 2007 Annual Compustat and I/B/E/S Summary files. The initial sample period covers the years 1988 to 2006. We perform the empirical analyses using observations from 1989 to 2005 because we need one year of lead and lag data. For inclusion in the final sample, each firm-year needs to have all variables required to estimate unexpected core earnings and at least 15 observations per industry-year. Because we use sales as a scalar, we also eliminate observations with sales less than $1 million to avoid outliers.

Table 1 presents a comparison between all Compustat firms and the firms included in our sample. Overall, 9 percent of the firms included in Compustat report discontinued operations in the test period. The percentage of firms reporting discontinued operations varies by year, with a low of 5 percent in 1994 to a high of 19 percent in 2006. The final sample used for our empirical analyses contains 79,643 firm-year observations, with 6,262 observations (8 percent) reporting discontinued operations. The pattern of annual percentages of firms reporting discontinued operations in the

TABLE 1

Sample Composition by Year

Fiscal Year	All Compustat Firms: Number of Firms	All Compustat Firms: Firms Reporting Discontinued Operations	All Compustat Firms: Percentage of Firms Reporting Discontinued Operations	Firms Used in Empirical Analyses: Number of Firms	Firms Used in Empirical Analyses: Firms Reporting Discontinued Operations	Firms Used in Empirical Analyses: Percentage of Firms Reporting Discontinued Operations	Firms Used in Empirical Analyses: Absolute Value of Discontinued Operations as a Percentage of Sales
1988	6318	584	9%				
1989	6240	607	10%	4,159	382	9%	5.1%
1990	6258	505	8%	4,487	351	8%	5.5%
1991	6340	481	8%	4,422	315	7%	5.5%
1992	6555	437	7%	4,153	239	6%	4.2%
1993	7074	477	7%	4,054	239	6%	3.9%
1994	8028	429	5%	4,885	265	5%	4.4%
1995	8326	494	6%	5,298	321	6%	5.8%
1996	9138	511	6%	5,512	342	6%	4.7%
1997	9227	546	6%	5,898	342	6%	5.9%
1998	8810	561	6%	5,705	340	6%	6.0%
1999	9028	647	7%	5,265	386	7%	5.8%
2000	8787	676	8%	5,160	366	7%	7.6%
2001	8261	667	8%	4,744	364	8%	7.5%
2002	7822	1006	13%	4,179	377	9%	5.3%
2003	7549	1082	14%	4,156	499	12%	3.5%
2004	7351	1166	16%	4,285	619	14%	2.8%
2005	7058	1247	18%	3,281	515	16%	2.8%
2006	5784	1089	19%				
Total	143,954	13,212	9%	79,643	6,262	8%	5.1%

Table 1 presents the data used in this study for the period 1988–2006. Since we need lead and lag data for one year, our empirical analyses use observations for the period 1989 through 2005. Columns 2–4 provide details of the total number of firms, and the number and percentage of firms reporting discontinued operations by year, respectively, in the Compustat database. Columns 5–7 show the number of firms with available data used in our empirical analyses, and the number and percentage of the firms reporting discontinued operations, respectively. The last column presents the absolute value of discontinued operations as a percentage of sales.

final sample is similar to that of the entire population. There is a significant increase in the percentage of firms reporting discontinued operations after SFAS No. 144 became effective, increasing from 8 percent in 2001 to 13 percent in 2002. Untabulated results show a significant difference between the mean percentage of firms reporting discontinued operations in the pre-SFAS No. 144 and post-SFAS No. 144 periods.[10]

We note a gradual decrease in the magnitude of discontinued operations in the post-SFAS No. 144 period, with the absolute value of discontinued operations decreasing from 7.5 percent in 2001 to 2.8 percent in 2006, consistent with the broader scope of SFAS No. 144. Untabulated results also show that the mean unexpected core earnings of firms reporting discontinued operations in the post-SFAS No. 144 period are lower than those in the pre-SFAS No. 144 period,[11] as we investigate in the following section.

The empirical analyses test whether firms increase core earnings by shifting operating expenses to discontinued operations. We follow McVay (2006) to measure core earnings, expected core earnings, unexpected core earnings, and the unexpected change in core earnings.[12] Core earnings is defined as operating earnings before depreciation and special items, scaled by revenue. To estimate unexpected core earnings and the unexpected change in core earnings, we use the following expectation models:

$$CE_t = \beta_0 + \beta_1 CE_{t-1} + \beta_2 ATO_t + \beta_3 ACCRUALS_{t-1} + \beta_4 ACCRUALS_t + \beta_5 \Delta SALES_t + \beta_6 NEG\_\Delta SALES_t + \varepsilon_t \quad (1)$$

$$\Delta CE_t = \varphi_0 + \varphi_1 CE_{t-1} + \varphi_2 \Delta CE_{t-1} + \varphi_3 \Delta ATO_t + \varphi_4 ACCRUALS_{t-1} + \varphi_5 ACCRUALS_t + \varphi_6 \Delta SALES_t + \varphi_7 NEG\_\Delta SALES_t + \nu_t, \quad (2)$$

where:

CE_t = core earnings, calculated as [Sales − Cost of Goods Sold − Selling, General and Administrative Expenses (data item #13)]/Sales (data item #12) for firms not reporting discontinued operations; for firms reporting discontinued operations we use restated data: Data (#117 − #131 − #132)/#117;

ΔCE_t = change in core earnings, calculated as $CE_t - CE_{t-1}$;

ATO_t = asset turnover ratio, defined as $Sales_t$ (# 12) / $(NOA_t + NOA_{t-1})$ / 2;

NOA = Net Operating Assets
= Operating Assets − Operating Liabilities
= [Total Assets (#6) − Cash (#1) and Short-Term Investments (#32)] − [Total Assets (#6) − Total Debt (#9 + #34) − Book Value of Common (#60) and Preferred Equity (#130) − Minority Interest (#38)];

[10] For Compustat firms, the mean reporting percentage increases from 7.03 percent in the pre-SFAS No. 144 period to 15.70 percent in the post-SFAS No. 144 period (t-stat = 25.12, $p < 0.01$). For our sample firms, the mean reporting percentage increased from 7.70 percent in the pre-SFAS No. 144 period to 16.05 percent in the post-SFAS No. 144 period (t-stat = 21.21, $p < 0.01$).

[11] For firms reporting discontinued operations, the mean unexpected core earnings in the pre- and post-SFAS No. 144 period are 3.74 percent and 0.92 percent, respectively (t-stat = 3.73, $p < 0.01$).

[12] Fan et al. (2010) observe that the association between unexpected core earnings and the magnitude of negative special items may be biased because noncash special items are included in accruals that are used as an independent variable in the expectation model. Since this study focuses on discontinued operations, the association between core earnings and discontinued operations is not affected by the design-driven bias. We also re-estimate the expectation models without using contemporaneous accruals and re-run our main analysis. The results are reported in the robustness section.

ΔATO_t = change in asset turnover, calculated as $ATO_t - ATO_{t-1}$
$ACCRUALS_t$ = operating accruals, calculated as [Net Income before Extraordinary Items (#123) − Cash from Operations (#308 − #124)]/Sales (#12);
$\Delta SALES_t$ = percentage change in sales, calculated as $[SALES_t(\#12) - SALES_{t-1}] / SALES_{t-1}$; and
$NEG\_\Delta SALES_t$ = percentage change in sales ($\Delta SALES_t$), if $\Delta SALES_t$ is less than zero, and 0 otherwise.

To provide evidence of classification shifting, we test for a positive association between unexpected core earnings and discontinued operations in year *t*. However, a positive association can arise mechanically due to the removal of a loss-making operation, because the expectation model for CE_t includes prior year core earnings (CE_{t-1}) as an independent variable. To address this possibility, we use restated data to calculate the variables included in the expectation models for firms with discontinued operations. The restated data exclude the operating income (or loss) of the discontinued operations from the reported operating income for year *t*−1. The Compustat database includes restated data for the previous ten years on the basis comparable to the current year for firms with discontinued operations, acquisitions or accounting changes.[13] We delete firm-year observations reporting acquisitions or accounting changes. We report the empirical results using restated data.

We measure the expected value of core earnings and the change in core earnings for firm *i* using the predicted values from Equations (1) and (2), respectively. We estimate each equation by industry-year, excluding firm *i* from the estimation. Unexpected core earnings and the unexpected change in core earnings are the difference between the actual and predicted values of Equations (1) and (2), respectively. The data on discontinued operations are collected from the Compustat annual database (item #66).

V. RESULTS

Descriptive Statistics

We multiply *%DO*, *%DO_NEG*, and *%DO_POS* by −1 to capture the positive association between discontinued operations and unexpected core earnings. Table 2 provides descriptive statistics for the variables used in the empirical analyses. Panel A presents the descriptive statistics for the entire sample. The mean (median) core earnings for all firms are 0.024 (0.103). The mean (median) annual sales of all firms in the full sample are $1,660M ($142M).

Table 2, Panel B presents the descriptive statistics for the subsample of firms reporting discontinued operations. Firms with discontinued operations are relatively larger in size, have higher average sales volume, lower sales growth, and lower core earnings than other firms. The mean (median) unexpected core earnings is 0.028 (0.011) for firms reporting discontinued operations, which is much higher than 0.002 (0.004) for the entire sample. This finding is consistent with our prediction that firms reporting discontinued operations shift recurring operating expenses to discontinued operations in order to increase core earnings.

In order to control for performance, we compare firm characteristics, such as firm size (*SIZE*), book-to-market ratio (*BM*), total accruals (*ACCRUALS*), operating cash flow (*OCF*), and return on assets (*ROA*), between firm-year observations with and without discontinued operations. We find

[13] To ensure the quality of the restated data, we randomly selected 50 observations reporting discontinued operations and checked for data consistency between the SEC filings and Compustat database. We find that the restated data in the Compustat database are consistent with the corresponding data in Form 10-K for the selected observations.

TABLE 2
Descriptive Statistics

Panel A: Descriptive Statistics for the Full Sample[a]

Variable	Mean	Median	Standard Deviation	25%	75%
$SALES_t$ (in millions)	1,660	142	8,180	29	679
$\Delta SALES_{t-1,t}$	0.209	0.091	0.572	−0.015	0.256
CE_t	0.024	0.103	0.572	0.029	0.195
$\Delta CE_{t-1,t}$	0.031	0.001	0.420	−0.028	0.030
$\Delta CE_{t,t+1}$	0.018	0.001	0.381	−0.030	0.028
UE_CE_t	0.002	0.004	0.248	−0.047	0.066
$UE\_\Delta CE_{t+1}$	0.003	0.004	0.219	−0.041	0.057
$\%DO_t$	0.001	0.000	0.015	0.000	0.000
$\%DO\_NEG_t$	0.002	0.000	0.013	0.000	0.000
$\%DO\_POS_t$	−0.001	0.000	0.007	0.000	0.000
SIZE	5.136	5.008	2.259	3.474	6.646
BM	0.627	0.505	1.144	0.271	0.855
$ACCRUALS_t$	−0.113	−0.051	0.316	−0.130	−0.004
OCF	0.053	0.074	0.177	0.005	0.138
ROA_t	−0.012	0.032	0.188	−0.029	0.075

Panel B: Descriptive Statistics for Firm-Years reporting Discontinued Operations[a]

Variable	Mean	Median	Standard Deviation	25%	75%
$SALES_t$ (in millions)	2,071	179	11,089	35	880
$\Delta SALES_{t-1,t}$	0.133	0.047	0.600	−0.084	0.187
CE_t	−0.008	0.081	0.572	0.011	0.151
$\Delta CE_{t-1,t}$	0.058	0.004	0.481	−0.026	0.046
$\Delta CE_{t,t+1}$	0.030	0.003	0.390	−0.025	0.033
UE_CE_t	0.028	0.011	0.298	−0.046	0.101

(continued on next page)

Panel B: Descriptive Statistics for Firm-Years reporting Discontinued Operations[a]

Variable	Mean	Median	Standard Deviation	25%	75%
$UE\_\Delta CE_{t+1}$	0.005	0.002	0.236	−0.044	0.057
$\%DO_t$	0.0133	0.002	0.053	−0.008	0.026
$\%DO\_NEG_t$	0.024	0.002	0.042	0.000	0.026
$\%DO\_POS_t$	−0.010	−0.000	0.023	−0.008	0.000
$SIZE$	5.347	5.237	2.266	3.610	6.961
BM	0.581	0.528	1.370	0.270	0.887
$ACCRUALS_t$	−0.123	−0.055	0.347	−0.137	−0.008
OCF	0.041	0.054	0.144	−0.003	0.111
ROA_t	−0.031	0.014	0.174	−0.054	0.052

[a] The maximum number of observations in Panels A and B are 79,643 and 6,262, respectively.

Variable Definitions:

$SALES_t$ = sales revenue in millions (Compustat data item #12);

$\Delta SALES_t$ = percentage change in sales, calculated as $[SALES_t\ (\#12) - SALES_{t-1}] / SALES_{t-1}$;

$\%DO_t$ = discontinued operations scaled by sales and multiplied by (−1); [Discontinued Operations (#66) × (−1)]/*SALES* (#12);

$\%DO\_NEG_t$ = discontinued operations (Compustat data item #66) scaled by sales and multiplied by (−1), when reported discontinued operations are income-decreasing, and 0 otherwise;

$\%DO\_POS_t$ = discontinued operations (Compustat data item #66) scaled by sales and multiplied by (−1), when reported discontinued operations are income-increasing, and 0 otherwise;

CE_t = core earnings (operating income before depreciation and special items), calculated as [Sales – Cost of Goods Sold – Selling, General and Administrative Expenses] (#13)]/ Sales (#12) for firms with no discontinued operations; for firms with discontinued operations, restated data are used including Data (#117 – #131 – #132)/#117;

ΔCE_t = change in core earnings, calculated as $CE_t - CE_{t-1}$;

UE_CE_t = unexpected core earnings measured as the difference between reported and predicted core earnings, where predicted values are estimated from the following equation:

$$CE_t = \beta_0 + \beta_1 CE_{t-1} + \beta_2 ATO + \beta_3 ACCRUALS_{t-1} + \beta_4 ACCRUALS_t + \beta_5 \Delta SALES_t + \beta_6 NEG\_\Delta SALES_t + \varepsilon_t; \quad (1)$$

$UE\_\Delta CE_{t+1}$ = unexpected change in core earnings is the difference between actual change in core earnings and predicted change in core earnings, where predicted values are estimated from the following equation:

$$\Delta CE_t = \varphi_0 + \varphi_1 CE_{t-1} + \varphi_1 \Delta CE_{t-1} + \varphi_3 \Delta ATO_t + \varphi_3 ACCRUALS_{t-1} + \varphi_4 ACCRUALS_t + \varphi_5 \Delta SALES_t + \varphi_6 NEG\_\Delta SALES_t + \nu_t; \quad (2)$$

(continued on next page)

ATO_t = asset turnover ratio, defined as $Sales_t$ (#12)/($NOA_t + NOA_{t-1}$) / 2:
NOA ≡ Net Operating Assets
≡ Operating Assets – Operating Liabilities
≡ [Total Assets (#6) – Cash (#1) and Short-Term Investment (#32)] – [Total Assets (#6) – Total Debt (#9 + #34) – Book Value of Common (#60) and Preferred Equity (#130) – Minority Interest (#38)];
ΔATO_t = change in asset turnover, calculated as $ATO_t - ATO_{t-1}$;
$NEG\_\Delta SALES_t$ = percentage change in sales ($\Delta SALES_t$), if $\Delta SALES_t$ is less than 0, and 0 otherwise;
$SIZE_t$ = natural logarithm of total assets;
BM_t = ratio of book value to market value;
$ACCRUALS_t$ = operating accruals, calculated as [Net Income before Extraordinary Items (#123) – Cash from Operation (#308 – # 124)]/Sales (#12);
OCF_t = operating cash flow scaled by lagged total assets; and
ROA_t = income before extraordinary items divided by average total assets.

significant differences in all these variables; consequently, we control for these variables in our empirical tests.[14]

Classification Shifting using Discontinued Operations

We follow McVay (2006) in designing our regression models to test whether firms increase core earnings by using classification shifting when reporting discontinued operations. We modify her equations by substituting discontinued operations (*%DO*) for special items (*%SI*) and adding control variables, as shown in the following equations:

$$UE\_CE_t = \alpha_0 + \alpha_1 \%DO_t + \alpha_2 SIZE_t + \alpha_3 BM_t + \alpha_4 ACCRUALS_t + \alpha_5 OCF_t + \alpha_6 ROA_t + \varepsilon_t \quad (3a)$$

$$UE\_\Delta CE_{t+1} = \eta_0 + \eta_1 \%DO_t + \eta_2 \%DO_{t+1} + \eta_3 SIZE_t + \eta_4 BM_t + \eta_5 ACCRUALS_t + \eta_6 OCF_t + \eta_7 ROA_t + \upsilon_{t+1}. \quad (3b)$$

UE_CE_t is unexpected core earnings in year t. $UE\_\Delta CE_{t+1}$ is the unexpected change in core earnings in year $t+1$. The variable of interest, $\%DO_t$, is the value of the aggregate results, losses, and gains from discontinued operations (scaled by sales). To test H1, we predict a positive association between unexpected core earnings and discontinued operations (i.e., $\alpha_1 > 0$). However, a positive association could also arise from an improvement in operations, due to the discontinuation of a loss-making operation. If this is the case, then the improvement in core earnings should persist in future years. However, if the improvement is due to classification shifting, we predict a negative association between the unexpected change in core earnings in year $t+1$ and the discontinued operations in year t (i.e., $\eta_1 < 0$). This reversal is due to operating expenses reappearing again as a part of core earnings. To control for discontinued operations in year $t+1$, we add $\%DO_{t+1}$ to Equation (3b). We predict a positive coefficient on $\%DO_{t+1}$ (i.e., $\eta_2 > 0$). Finally, we add the following control variables to Equations (3a) and (3b): firm size ($SIZE_t$), book-to-market ratio (BM_t), accruals ($ACCRUALS_t$), operating cash flow (OCF_t) and return on assets (ROA_t). We make no prediction as to the sign of these control variables.

A firm's classification shifting behavior could vary depending on whether firms report overall gains or losses from discontinued operations. For example, managers might have more incentive or greater discretion to shift more operating expenses when discontinued operations already have losses.[15] To test this conjecture, we replace the discontinued operations variable *%DO* with two variables, *%DO_NEG* for discontinued operations with overall losses and *%DO_POS* for discontinued operations with overall gains, in the following equations:

$$UE\_CE_t = \pi_0 + \pi_1 \%DO\_NEG_t + \pi_2 \%DO\_POS_t + \pi_3 SIZE_t + \pi_4 BM_t + \pi_5 ACCRUALS_t + \pi_6 OCF_t + \pi_7 ROA_t + \psi_t; \quad (4a)$$

$$UE\_\Delta CE_{t+1} = \varphi_0 + \varphi_1 \%DO\_NEG_t + \varphi_2 \%DO\_POS_t + \varphi_3 \%DO\_NEG_{t+1} + \varphi_4 \%DO\_POS_{t+1} + \varphi_5 SIZE_t + \varphi_6 BM_t + \varphi_7 ACCRUALS_t + \varphi_8 OCF_t + \varphi_9 ROA_t + \omega_{t+1}. \quad (4b)$$

If firms engage in classification shifting using discontinued operations irrespective of overall gains or losses, we predict both π_1 and π_2 (φ_1 and φ_2) to be positive (negative).

---

[14] The differences between the means of the variables *SIZE*, *BM*, *ACCRUALS*, and *ROA* are significant at better than the 1 percent level. The difference between the means of the variable *OCF* is significant at the 10 percent level.

[15] For example, a manager may have more of an incentive to blame poor past results of the firm as a whole, on discontinued operations sold at a loss. One way of doing this is to increase the operating loss of the discontinued operations. In addition, the big bath hypothesis predicts that companies may deliberately take large losses in a single year through write-offs and write-downs in order to increase accounting flexibility in future years.

TABLE 3

Regressions of Unexpected Core Earnings and Unexpected Change in Core Earnings on Discontinued Operations

Panel A: Dependent Variable = UE_CE_t

Independent Variables	Coefficient	t-stat	Coefficient	t-stat
Intercept	−0.022	−8.70***	−0.023	−9.13***
$\%DO_t$	0.451	7.73***		
$\%DO\_NEG_t$			0.705	10.68***
$\%DO\_POS_t$			−0.480	−3.76***
$SIZE_t$	0.000	−0.32	0.000	−0.43
BM_t	0.000	−0.24	0.000	−0.14
$ACCRUALS_t$	−0.036	−3.16***	−0.037	−3.24***
OCF_t	0.373	34.72***	0.373	34.75***
ROA_t	−0.015	−1.21	−0.012	−0.99
Adjusted R^2	7.0%		7.1%	

Panel B: Dependent Variable = $UE\_\Delta CE_{t+1}$

Independent Variables	Coefficient	t-stat	Coefficient	t-stat
Intercept	−0.027	−11.65***	−0.027	−11.92***
$\%DO_t$	−0.124	−2.31**		
$\%DO_{t+1}$	0.419	8.28***		
$\%DO\_NEG_t$			−0.164	−2.67***
$\%DO\_POS_t$			−0.090	−0.76
$\%DO\_NEG_{t+1}$			0.650	10.97***
$\%DO\_POS_{t+1}$			−0.245	−2.39***
$SIZE_t$	0.004	10.83***	0.004	10.66***
BM_t	0.004	5.46***	0.004	5.50***
$ACCRUALS_t$	0.104	9.93***	0.105	9.97***
OCF_t	0.164	16.71***	0.164	16.79***
ROA_t	−0.132	−12.06***	−0.131	−11.93***
Adjusted R^2	0.75%		0.82%	
No. of Observations	72,372		72,372	

*, **, *** Indicate significance at the 10, 5, and 1 percent levels, respectively. All the test results use a two-tailed t-test except $\%DO_t$, $\%DO\_NEG_t$, $\%DO\_POS_t$, $\%DO_{t+1}$, $\%DO\_NEG_{t+1}$, and $\%DO\_POS_{t+1}$ (use a one-tailed t-test).

Discontinued operations (*%DO*) are scaled by sales multiplied by (−1); [Discontinued Operations (#66) × (−1)] /SALES (#12). Discontinued operations (Compustat data item #66) scaled by sales and multiplied by (−1), when reported discontinued operations are income-decreasing, and 0 otherwise. Discontinued operations (Compustat data item #66) are scaled by sales and multiplied by (−1) when reported discontinued operations are income-increasing, and 0 otherwise. All other variables are as defined in Table 2.

Table 3, Panel A reports regression results for Equations (3a) and (4a). For Equation (3a), the coefficient on $\%DO_t$ is both positive and significant ($\alpha_1 = 0.451$, $p < 0.01$), a result consistent with H1.[16] For Equation (4a), the coefficient on $\%DO\_NEG_t$ is both positive and significant

[16] All the test results reported in this section are based on a one-tailed test of significance, except where otherwise noted.

($\pi_1 = 0.705$, $p < 0.01$); however the coefficient on $\%DO\_POS_t$ is negative and significant ($\pi_2 = -0.480$, $p < 0.01$), opposite of the predicted sign. These findings suggest that classification shifting takes place when firms report losses from discontinued operations. Economically, one standard deviation increase in the loss from discontinued operations is expected to increase unexpected core earnings (scaled by sales) by approximately 92 basis points (0.705 [the coefficient on $\%DO\_NEG_t$ from Table 3] multiplied by 0.013 [the standard deviation of income-decreasing discontinued operations for this sample from Table 2]).

Table 3, Panel B reports regression results of Equation (3a) and (4b). For Equation (3b), the coefficient on $\%DO_t$ is both negative and significant ($\eta_1 = -0.124$, $p < 0.01$), a result consistent with operating expenses reappearing as a component of core earnings. For Equation (4b), only the coefficient on $\%DO\_NEG_t$ is both negative and significant ($\varphi_1 = -0.164$, $p < 0.01$). The coefficient on $\%DO\_POS_t$ is negative and insignificant ($\varphi_2 = -0.090$, $p = 0.45$). These findings suggest the reversals of the improvements in core earnings are associated with the reporting of losses from discontinued operations in year t. Consistent with Panel A, there is a contemporaneous positive association between $UE\_\Delta CE_{t+1}$ and each of the following variables $\%DO_{t+1}$ ($\eta_2 = 0.419$, $p < 0.01$) and $\%DO\_NEG_{t+1}$ ($\varphi_3 = 0.650$, $p < 0.01$). Consistent with Panel A, but inconsistent with our prediction $\%DO\_POS_{t+1}$ is significantly negative ($\varphi_4 = -0.245$, $p < 0.01$).

Overall, findings in Table 3 provide evidence consistent with classification shifting using discontinued operations. However, our predictions only hold for firms reporting income-decreasing discontinued operations. We do not find consistent evidence for firms reporting income-increasing discontinued operations.[17] Hence, we use $\%DO\_NEG_t$ as our variable of interest in all subsequent analyses.

Meeting or Beating Earnings Benchmarks

We use the following equations to examine whether classification shifting using discontinued operations is used to meet or beat certain earnings benchmarks:

$$UE\_CE_t = \gamma_0 + \gamma_1 \%DO\_NEG_t + \gamma_2 MEET\&BEAT_t + \gamma_3 MEET\&BEAT_t * \%DO\_NEG_t + \gamma_4 SIZE_t + \gamma_5 BM_t + \gamma_6 ACCRUALS_t + \gamma_7 OCF_t + \gamma_8 ROA_t + \xi_t \quad (5a)$$

$$UE\_\Delta CE_{t+1} = \delta_0 + \delta_1 \%DO\_NEG_t + \delta_2 MEET\&BEAT_t + \delta_3 MEET\&BEAT_t * \%DO\_NEG_t + \delta_4 \%DO\_NEG_{t+1} + \delta_5 SIZE_t + \delta_6 BM_t + \delta_7 ACCRUALS_t + \delta_8 OCF_t + \delta_9 ROA_t + \xi_{t+1}. \quad (5b)$$

$MEET\&BEAT_t$ is an indicator variable taking on a value of 1 if the firm-year observation meets or beats one of the following three earnings benchmarks, and 0 otherwise. The first earnings benchmark is zero earnings. This variable proxies for earnings management to avoid losses, where $MEET\&BEAT_t = 1$ if the observation has earnings per share greater than or equal to zero, and 0 otherwise. The second earnings benchmark is prior year's earnings. This variable proxies for earnings management to avoid a decline in earnings from the previous year, where $MEET\&BEAT_t = 1$ if the observation has earnings per share changes greater than or equal to zero. The earnings subtotal used for benchmarks 1 and 2 are earnings before discontinued operations and extraordinary items because they have been widely used as performance measures in accounting literature. The third benchmark is analysts' forecasts. Following Matsumoto (2002),

[17] We believe that classification shifting can take place when firms report positive discontinued operations, although we do not find the predicted results. We leave this issue to future research.

$MEET\&BEAT_t = 1$ if firm-years have positive analysts' forecast errors,[18] and 0 otherwise. We use the median consensus forecast available before the earnings announcement as obtained from the adjusted I/B/E/S summary database.

Table 4, Panel A reports the regression results for Equation (5a). The coefficients on $\%DO\_NEG_t$ are consistently positive and significant for benchmarks 1 and 2. The coefficient on benchmark 3 is positive but insignificant. The coefficients on the variable of interest, $MEET\&BEAT_t$ interacted with $\%DO\_NEG_t$ are positive and significant. These findings are consistent with H2a, H2b, and H2c. Table 4, Panel B reports the regression results for Equation (5b). The coefficient on $\%DO\_NEG_t$ is negative and significant only for benchmark 1. The coefficient on benchmark 2 is negative but insignificant. The coefficient on benchmark 3 is positive and significant. The coefficient on the variable of interest, $MEET\&BEAT_t$ interacted with $\%DO\_NEG_t$, is only significant for benchmark 3. These findings are consistent with H2c.

Overall, the results suggest that companies reporting income-decreasing discontinued operations use classification shifting to increase core earnings in order to meet or beat the analysts' forecast benchmark.

Effect of SFAS No. 144

As shown in Table 1, the recognition of discontinued operations significantly increased after SFAS No. 144 became effective in 2002. This section further examines the impact of the change in accounting standards on the association between discontinued operations and unexpected core earnings presented in Table 3. We modify Equations (5a) and (5b) by replacing $MEET\&BEAT_t$, with *POST_2002*, which takes a value of 1 if the discontinued operations occur in fiscal year 2002 or later, and 0 otherwise. Table 5 presents the results for both of these models.

The second and third columns of Table 5 report the results of the modified Equation (5a). Consistent with the findings reported in previous tables, the discontinued operations variable, $\%DO\_NEG_t$ (0.774, $p < 0.01$), is positively associated with unexpected core earnings. The coefficient on the interaction term is both negative and significant (-0.411, $p < 0.01$). The fourth and fifth columns of Table 5 report the results of the modified Equation (5b). Consistent with the findings reported in previous tables, the discontinued operations variable, $\%DO\_NEG_t$ (-0.211, $p < 0.01$), is negatively associated with the $t+1$ change in unexpected core earnings. The coefficient on the interaction term is both positive and significant (0.329, $p = 0.02$). Taken together, these results suggest a reduction in the magnitude of classification shifting using discontinued operations in the post-SFAS No. 144 period.

There are two potential reasons for this finding. First, SFAS No. 144 broadened the definition of discontinued operations by replacing the "business segment" requirement under APB No. 30 with the "component of an entity concept." This change allowed smaller asset dispositions to be reported as discontinued operations, increasing the reporting frequency. In addition, allowing smaller asset dispositions to qualify as discontinued operations potentially reduces the amount of operating expenses that can be shifted to the disposition. Second, the Sarbanes-Oxley Act was enacted around the same time as SFAS No. 144, and may have led to a decrease in classification shifting due to increased regulation and oversight.

Robustness Tests

To investigate whether classification shifting using income-decreasing discontinued operations is concentrated among firms frequently reporting discontinued operations, we perform the following analysis. Table 6, Panel A reports that approximately 26 percent (3,105/11,895) of our

[18] Forecast errors are defined as actual earnings per share less the last-issued median consensus analyst forecast.

TABLE 4

Regressions of Unexpected Core Earnings and Unexpected Changes in Core Earnings Earnings Benchmarks Analyses

Panel A: Dependent Variable = UE_CE_t

	Benchmark 1		Benchmark 2		Benchmark 3	
	Zero Earnings		Prior Period Earnings		Analysts' Forecasts	
Independent Variables	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	−0.015	−5.19***	−0.033	−11.78***	−0.056	−14.80***
$\%DO\_NEG_t$	0.552	6.85***	0.438	4.70***	0.047	0.39
$MEET\&BEAT_t$	−0.012	−4.56***	0.015	7.95***	−0.006	−2.81***
$MEET\&BEAT_t * \%DO\_NEG_t$	0.407	2.89***	0.525	3.99***	1.133	5.76***
$SIZE_t$	0.000	0.29	0.000	0.29	0.005	8.40***
BM_t	0.000	−0.48	0.000	0.49	0.005	3.26***
$ACCRUALS_t$	−0.034	−2.94***	−0.041	−3.53***	−0.039	−2.75***
OCF_t	0.375	34.94***	0.373	34.80***	0.377	27.95***
ROA	0.002	0.15	−0.022	−1.80*	−0.034	−2.20**
Adjusted R^2	7.06%		7.14%		8.0%	

Panel B: Dependent Variable = $UE\_\Delta CE_{t+1}$

	Benchmark 1		Benchmark 2		Benchmark 3	
	Zero Earnings		Prior Period Earnings		Analysts' Forecasts	
Independent Variables	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	−0.028	−10.42***	−0.026	−9.99***	−0.035	−10.21***
$\%DO\_NEG_t$	−0.158	−2.13**	−0.095	−1.11	0.342	3.11***
$MEET\&BEAT_t$	0.002	0.78	−0.002	−1.22	−0.002	−1.14
$MEET\&BEAT_t * \%DO\_NEG_t$	0.025	0.20	−0.114	−0.94	−0.476	−2.68***
$\%DO\_NEG_{t+1}$	0.649	10.95***	0.646	10.91***	0.527	6.56***
$SIZE_t$	0.004	10.59***	0.004	10.66***	0.005	10.02***

(continued on next page)

Panel B: Dependent Variable = $UE\_\Delta CE_{t+1}$

	Benchmark 1		Benchmark 2		Benchmark 3	
	Zero Earnings		Prior Period Earnings		Analysts' Forecasts	
Independent Variables	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
BM_t	0.004	5.50***	0.004	5.34***	0.006	4.72***
$ACCRUALS_t$	0.104	9.93***	0.105	10.03***	0.135	10.53***
OCF_t	0.164	16.73***	0.164	16.77***	0.210	17.24***
ROA	−0.134	−11.73***	−0.130	−11.81***	−0.164	−11.64***
Adjusted R^2	0.81%		0.81%		1.24%	

*, **, *** Indicate significance at the 10, 5, and 1 percent levels, respectively. All the test results use a two-tailed t-test except $\%DO\_NEG_t$, $MEET\&BEAT_t * \%DO\_NEG_t$, and $\%DO\_NEG_{t+1}$ (use a one-tailed t-test).

$MEET\&BEAT_t$ is an indicator variable that takes a value of 1 if the firm-year observation meets or exceeds an earnings benchmark, and 0 otherwise.

Benchmark 1 Zero Earnings: where $MEET\&BEAT_t = 1$ if the earnings per share of a firm (Compustat annual data item #58) is non-negative, and 0 otherwise.
Benchmark 2 Prior Period Earnings: earnings management to avoid an earnings decline, where $MEET\&BEAT_t = 1$ if $EPS_t - EPS_{t-1}$ is non-negative, and 0 otherwise.

Benchmark 3 Analysts' Forecasts: where $MEET\&BEAT_t = 1$ if firm-years have analysts' forecast errors [Actual EPS–Consensus Forecast] that are non-negative, and 0 otherwise.

Number of observations that meet or exceeds:

(1) Benchmark 1: 50,828 out of 72,372 observations;

(2) Benchmark 2: 41,659 out of 72,372 observations; and

(3) Benchmark 3: 24,272 out of 41,720 observations.

All other variables are as defined in Table 2.

TABLE 5

Regressions of Unexpected Core Earnings and Unexpected Change in Core Earnings Pre and Post SFAS No. 144 Period

	Dependent Variables			
	UE_CE_t		$UE\_\Delta CE_{t+1}$	
Independent Variables	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	−0.022	−8.91***	−0.026	−11.50***
$\%DO\_NEG_t$	0.774	10.63***	−0.211	−3.14***
POST_2002	−0.005	−2.18**	−0.008	−3.92***
$\%DO\_NEG_t$ * *POST_2002*	−0.411	−2.38***	0.329	2.09**
$\%DO\_NEG_{t+1}$			0.646	10.91***
$SIZE_t$	0.000	0.03	0.005	11.14***
BM_t	0.000	−0.23	0.004	5.30***
$ACCRUALS_t$	−0.036	−3.17***	0.105	10.02***
OCF_t	0.373	34.76***	0.166	16.89***
ROA_t	−0.013	−1.12	−0.132	−12.08***
Adjusted R^2	7.1%		0.84%	
No. of Observations	72,372		72,372	

*, **, *** Indicate significance at the 10, 5, and 1 percent levels, respectively. All the test results use a two-tailed t-test except $\%DO\_NEG_t$, $\%DO\_NEG_t$ * *POST_2002*, and $\%DO\_NEG_{t+1}$ (use a one-tailed t-test).

POST_ 2002 is an indicator variable that takes a value of 1 if an observation is from fiscal year 2002 and afterwards, and 0 otherwise.

All other variables are as defined in Table 2.

sample firms have reported discontinued operations at least once during the sample period. However, only 1.6 percent (191/11,895) of our sample companies recognized discontinued operations more than four times during the sample period. Consequently, we do not believe that classification shifting using income-decreasing discontinued operations is concentrated among serial recognizers of discontinued operations. In order to confirm this belief, we re-estimate Equations (3a) and (3b) (replacing $\%DO_t$ with $\%DO\_NEG_t$) on a sample of firms that recognized discontinued operations only once during the sample period. Table 6, Panel B reports the results of the re-estimation. Overall, the results are consistent with the results reported in Table 3.

As a further robustness check, we re-estimated Equations (3a) and (3b), (replacing $\%DO_t$ with $\%DO\_NEG_t$) after deleting those companies recognizing more than four discontinued operations during the sample period. These results (untabulated) are also consistent with the main results. Overall, our results are consistent with classification shifting taking place when a firm reports income-decreasing discontinued operations.

To control for the effect of special items, we repeat all our analyses adding a continuous variable for income-decreasing special items (scaled by sales) as an independent variable in Equations (3a), (3b), (4a), and (4b). The results (not tabulated) are similar both qualitatively and quantitatively to what is reported in our tables.

Fan et al. (2010) observe that the classification shifting results documented in McVay (2006) could be subject to a research design problem. They assert that the associations between income-decreasing special items and unexpected core earnings and unexpected change in future core

TABLE 6
Additional Tests

Panel A: Frequency of Reporting Discontinued Operations

Frequency of Reporting Discontinued Operations	Number of Firms	Firm-Year Observations
0	8790	52,018
1	1516	11,767
2	794	7,318
3	407	3,993
4	197	2,177
More than 4 times	191	2,370

Panel B: Regression for Firms with One-Time Recognition of Discontinued Operations

Independent Variables	Dependent Variable = UE_CE_t		Dependent Variable = $UE\_\Delta CE_{t+1}$	
	Coefficient	t-stat	Coefficient	t-stat
Intercept	−0.026	−1.96**	−0.012	−1.02
$\%DO\_NEG_t$	0.412	2.71***	−0.297	−2.20**
$SIZE_t$	0.001	0.30	0.003	1.64*
BM_t	0.009	2.25**	0.012	3.29***
$ACCRUALS_t$	−0.201	−3.50***	0.008	0.16
OCF_t	0.284	5.26***	−0.035	−0.72
ROA_t	0.035	0.61	0.010	0.18
Adjusted R^2	5.2%		0.47%	

*, **, *** Indicate significance at the 10, 5, and 1 percent levels, respectively. All the test results use a two-tailed t-test except $\%DO\_NEG_t$ (uses a one-tailed t-test).

The sample consists of 1,516 observations representing firms that report discontinued operations only once (year t) during the sample period along with year t–1 and t+1 observations. All the variables are as defined previously.

earnings may be biased because noncash special items are included in total accruals that are used as an independent variable in the expectation model. We do not believe our study suffers from this potential bias because the results from discontinued operations are reported separately from the results of continuing operations, which are used to estimate unexpected core earnings and accruals. Nevertheless, as a robustness check, we remove contemporaneous accruals from the expectation models and re-estimate unexpected core earnings and unexpected changes in core earnings. We then re-estimate Equations (3a), (3b), (4a), and (4b) reported in Table 3. The results (untabulated) are qualitatively similar to those in Table 3. For Equations (3a) and (4a), the coefficients on $\%DO_t$ ($\alpha_1 = 0.436$, $p < 0.01$) and $\%DO\_NEG_t$ ($\pi_1 = 0.752$, $p < 0.01$) are both positive and significant. For Equations (3b) and (4b), the coefficients on $\%DO_t$ ($\eta_1 = -0.080$, $p = 0.06$) and $\%DO\_NEG_t$ ($\varphi_1 = -0.103$, $p = 0.05$) are both negative and significant. Overall, these results suggest that our findings are robust to the potential model specification error observed by Fan et al. (2010).

VI. CONCLUSIONS

This study examines whether managers use discontinued operations to increase core earnings by engaging in classification shifting. Following a methodology similar to that employed by McVay (2006), we find a positive association between unexpected core earnings and discontinued operations. Our analysis shows that the positive and significant association between unexpected core earnings and discontinued operations is driven by firms with losses from discontinued operations. We interpret these results as evidence that firms engage in classification shifting using discontinued operations. We do not find consistent evidence for firms reporting income-increasing discontinued operations, although classification shifting can also take place in those situations. We leave further investigation of this issue to future research. We also investigate the motivations to manage earnings using discontinued operations, finding that firms reporting income-decreasing discontinued operations use classification shifting to meet or beat analyst forecasts.

Our results also show that the magnitude of classification shifting using discontinued operations has declined since the introduction of SFAS No. 144. Two explanations potentially account for this finding. First, SFAS No. 144 broadened the definition of discontinued operations by replacing the business segment requirement under APB No. 30 with the component of an entity concept. This change allowed firms to report smaller asset dispositions as discontinued operations, increasing the reporting frequency. In addition, allowing smaller asset dispositions to qualify as discontinued operations potentially reduced the amount of operating expenses that could be shifted to the disposition. Second, the Sarbanes-Oxley Act was enacted around the same time as SFAS No. 144, and could have led to a decrease in classification shifting due to increased regulation and oversight.

Our results are robust to examining a sample of firms reporting discontinued operations only once during our sample period and controlling for income-decreasing special items. We also re-estimate the expectation models without using contemporaneous accruals and find consistent results.

We believe the findings of this study could have implications for the convergence project on discontinued operations between the FASB and the IASB. We also believe the study would benefit investors by alerting them to the potential use of income-decreasing discontinued operations for earnings management purposes.

REFERENCES

Accounting Principles Board (APB). 1973. *Reporting the Results of Operations*. No. 30. New York, NY: AICPA.

Barnea, A., J. Ronen, and S. Sadan. 1976. Classificatory smoothing of income with extraordinary items. *The Accounting Review* 51 (1): 110–122.

Barth, M., J. Elliot, and M. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (2): 387–413.

Bradshaw, M., and R. Sloan. 2002. GAAP versus the street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (1): 41–66.

Brown, L., and M. Caylor. 2005. A temporal analysis of thresholds: Propensities and valuation consequences. *The Accounting Review* 80 (2): 423–440.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (1): 99–126.

Bushee, B. 1998. The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* 73 (3): 305–333.

Cohen, D., A. Dey, and T. Lys. 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods. *The Accounting Review* 83 (3): 757–787.

Dechow, P. M., and D. J. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (2): 235–250.

———, R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, and ———. 1991. Executive incentive and horizon problem. *Journal of Accounting and Economics* 14 (1): 51–89.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *The Journal of Business* 72 (1): 1–33.

Fan, Y., A. Barua, W. M. Cready, and W. Thomas. 2010. Managing earnings using classification shifting: Evidence from quarterly special items. *The Accounting Review* 85 (4): 1303–1323.

Financial Accounting Standards Board (FASB). 2001a. *Accounting for the Impairment or Disposal of Long-Lived Assets*. SFAS No. 144. Financial Accounting Series. Norwalk, CT: FASB.

———. 2001b. *Goodwill and Other Intangible Assets*. SFAS No. 142. Financial Accounting Series. Norwalk, CT: FASB.

———. 1997. *Disclosures about Segments of an Enterprise and Related Information*. SFAS No. 131. Financial Accounting Series. Norwalk, CT: FASB.

———. 2004. *Applying the Conditions in Paragraph 42 of FASB Statement No. 144, Accounting for the Impairment or Disposal of Long-Lived Assets, in Determining Whether to Report Discontinued Operations*. Issue No. 03–13. Emerging Issues Task Force. Norwalk, CT: FASB.

———. 2001c. *Accounting for the Impact of the Terrorist Attacks of September 11, 2001*. Issue No. 01–10. Emerging Issues Task Force. Norwalk, CT: FASB.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (2): 125–153.

Healy, P. M., and J. M. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (4): 365–383.

International Accounting Standards Board (IASB). 2004. *Non-Current Assets Held for Sale and Discontinued Operations*. Standard 5. London, U.K.: IASB.

Jiang, J. 2008. Beating earnings benchmarks and the cost of debt. *The Accounting Review* 83 (2): 377–416.

Kinney, M., and R. Trezevant. 1997. The use of special items to manage earnings and perceptions. *Journal of Financial Statement Analysis* (1): 45–53.

Kothari, S. P. 2001. Capital market research in accounting. *Journal of Accounting and Economics* 31 (1–3): 105–231.

Lipe, R. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (Supplement): 37–64.

Lobo, G., and J. Zhou. 2006. Did conservatism in financial reporting increase after the Sarbanes-Oxley Act? Initial evidence. *Accounting Horizons* 20 (1): 57–73.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

Matsunaga, S., and C. Park. 2001. The effect of missing a quarterly earnings benchmark on the CEO's annual bonus. *The Accounting Review* 76 (3): 313–332.

McVay, S. 2006. Earnings management using classification shifting: An examination of core earnings and special items. *The Accounting Review* 81 (3): 501–532.

Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Payne, L. J., and S. G. Robb. 2000. Earnings management: The effect of *ex ante* earnings expectations. *Journal of Accounting, Auditing & Finance* 15 (4): 371–392.

Riedl, E., and S. Srinivasan. 2010. Signaling firm performance through financial statement presentation: An analysis using special items. *Contemporary Accounting Research* 27 (1): 289–332.

Ronen, J., and S. Sadan. 1975. Classificatory smoothing: Alternative income models. *Journal of Accounting Research* 13 (1): 133–149.

Roychowdhury, S. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42 (3): 335–370.

Xu, R., G. K. Taylor, and M. T. Dugan. 2007. Review of real earnings management literature. *Journal of Accounting Literature* 46: 195–228.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1511–1543

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1511

Determinants of CEO Pay: A Comparison of ExecuComp and Non-ExecuComp Firms

Brian Cadman
The University of Utah

Sandy Klasa
The University of Arizona

Steve Matsunaga
University of Oregon

ABSTRACT: We document that firms included in the ExecuComp database tend to be larger, more complex, followed by more analysts, have greater stock liquidity levels, and have higher total, but less concentrated, institutional ownership than other firms. Based on these differences, we test and find support for three predictions. First, ExecuComp firms rely more heavily on earnings and stock returns in determining CEO cash compensation. Second, the weight on earnings is more sensitive to differences in the extent of growth opportunities for ExecuComp firms. Third, the positive relation between institutional ownership concentration and the value of stock option grants is stronger for ExecuComp firms. Overall, our results suggest that ExecuComp and non-ExecuComp firms operate in different contracting environments that lead to differences in the design of their executive compensation contracts. As a result, care should be taken in extending results based on ExecuComp samples to non-ExecuComp firms.

Keywords: *ExecuComp database; executive compensation; stock option grants; contracting environment.*

Data Availability: *All data used in this study are publicly available from sources identified in the text.*

We gratefully acknowledge the comments of Steven Kachelmeier (senior editor), Steven Huddart (editor), two anonymous reviewers, Tom Bates, Mary Billings, John Campbell, Mary Ellen Carter, Lei Chen, Dave Guenther, David Haushalter, Kathy Kahle, Eric Kelley, Tom Lys, Sarah McVay, Wayne Mikkelson, Dale Morse, Terry Shevlin, Mike Stegemoller, and workshop participants at The University of Arizona, University of Illinois at Chicago, and University of Oregon, as well as participants at the 2006 AAA Annual Meeting. In addition, we recognize the research assistance provided by Richard Bilbas, Adeel Javed, Wing-bun Ng, and Scott Savoian.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Steven Huddart for serving as editor on previous versions.

Submitted: October 2006
Accepted: December 2009
Published Online: August 2010

I. INTRODUCTION

In recent years, researchers have relied extensively on Standard & Poor's (S&P) ExecuComp database, which covers firms in the S&P 500, S&P 400 mid-cap, and S&P 600 small-cap indices, to investigate a wide variety of corporate governance and executive compensation issues.[1] The database includes most of the large publicly traded firms in the United States, as well as medium and smaller size firms that provide a degree of heterogeneity. Although the firms in ExecuComp comprise approximately 88 percent of the market capitalization of publicly traded firms in the United States, they only represent 25 percent of the firms in the Compustat database. Thus, firms excluded from the ExecuComp database comprise a substantial portion of samples for empirical studies that do not require compensation data. Yet, little is known about their compensation structures or how the determinants of their CEO incentive contracts differ from those of ExecuComp firms. In this study, we document differences between ExecuComp and non-ExecuComp firms across multiple dimensions that are likely to affect a firm's contracting environment and provide evidence regarding how compensation functions (determinants of CEO cash compensation and option grants) differ across the two sets of firms.

Because the ExecuComp database only includes firms from the three major S&P indices, the firms in the database have a common and distinct set of characteristics that are likely to impact the firm's contracting environment. For example, ExecuComp firms tend to be more complex, with operations in more business and geographic segments, and higher values for Herfindahl-Hirschman concentration indices measuring business and geographic diversification. On average, ExecuComp firms also have a greater degree of analyst following and stock liquidity. In addition, ExecuComp firms have a larger proportion of institutional ownership and less concentrated institutional ownership (i.e., ownership is distributed across a greater number of institutions). Collectively, we expect these differences to influence the contracting environment and, as a result, the use and type of performance metrics in contract design. To provide evidence on how the differences in the contracting environment can lead to systematic differences in the determinants of CEO compensation, we test three empirical predictions.

Our first prediction is that firms in the ExecuComp database rely more heavily on earnings and stock returns as performance measures in determining CEO cash compensation. As shown by Holmstrom (1979) and Banker and Datar (1989), the weight placed on a given performance measure in an optimal incentive contract is proportional to the signal-to-noise ratio of the measure (i.e., the extent to which it reflects the agent's actions, as opposed to external factors beyond the agent's control). In our setting, we expect the relatively high level of analyst following associated with ExecuComp firms to reflect greater public scrutiny over reported earnings, thereby leading to a higher degree of earnings quality and lower noise in earnings as a performance measure. In addition, we expect the higher degree of stock liquidity for ExecuComp firms, combined with the higher degree of analyst following, to increase the efficiency with which market prices impound the manager's actions into firm value, thereby reducing the noise in stock returns as a measure of performance. Finally, we expect the costs of identifying, defining, and weighting individual performance measures to be greater for the more complex, diverse firms that are included in the ExecuComp database, thereby leading ExecuComp firms to rely more heavily on aggregate financial performance measures. The combination of having less noise in earnings and stock returns,

[1] From 1999–2006, we found 28 articles published in *The Accounting Review*, *Journal of Accounting and Economics*, and the *Journal of Accounting Research*, and 34 articles published in the *Journal of Financial Economics*, *Journal of Finance*, and the *Review of Financial Studies*, that used the ExecuComp database. Although we restricted our search to accounting and finance journals, papers that investigate corporate governance and executive compensation issues published in journals in other fields, such as economics and management, are also likely to rely on ExecuComp data.

along with a higher cost in defining individual measures, should lead ExecuComp firms to place a greater weight on both earnings and stock returns in determining CEO compensation.

We test this prediction by using the standard empirical model that relates cash compensation to accounting and stock return measures of performance (i.e., Lambert and Larcker 1987; Sloan 1993; Baber et al. 1996; Baber et al. 1999; Matsunaga and Park 2001). Consistent with our prediction, we find the coefficients on both accounting and stock returns are significantly larger for ExecuComp firms. These results are robust to the exclusion of loss firms, controlling for differences in CEO ownership, and using a predicted classification based on characteristics associated with the firm's contracting environment.[2] The latter result provides evidence that our findings reflect differences in the firm's contracting environment, as opposed to the classification as an ExecuComp firm.

While our first prediction relates to differences in the average weight on earnings and stock returns between the ExecuComp and non-ExecuComp samples, our second prediction relates to differences in the weight on earnings across firms within each sample. Baber et al. (1996) and Ittner et al. (1997) note that the conservatism in the accounting system reduces the ability of accounting earnings to properly reflect CEO performance for growth firms, positing that the relative weight placed on earnings is negatively related to the firm's investment opportunity set (IOS). Their theory implicitly assumes that, in the absence of growth opportunities, firms use earnings as a primary performance measure. While this assumption may be appropriate for ExecuComp firms, it is less likely to be valid for non-ExecuComp firms. As noted above, independent of growth opportunities, we expect non-ExecuComp firms to substitute individual, nonfinancial measures for earnings in their CEO bonus plans. Thus, while high-growth and low-growth non-ExecuComp firms are likely to choose different individual performance measures, we would expect both high- and low-growth non-ExecuComp firms to place similarly low weights on earnings. This leads to our prediction that the weight on earnings should be less sensitive to the firm's IOS for non-ExecuComp firms than for ExecuComp firms.[3]

Supporting this prediction, we find the predicted negative association between the extent of a firm's investment opportunity set and the weight placed on earnings in our sample of ExecuComp firms, but not in our sample of non-ExecuComp firms. In addition, we document that the shift in weights from earnings toward stock returns as the IOS increases is significantly larger for ExecuComp firms. This result is robust to adjustments for loss firms, CEO ownership, and using predicted classifications.

Although our first two predictions relate to cash compensation, the differences in the contracting environment also have implications for the determinants of CEO stock option grant incentives. Specifically, we predict that the positive association between institutional ownership concentration and the extent of incentives in the CEO's annual stock option grant documented by Hartzell and Starks (2003) for ExecuComp firms will be less likely to hold for non-ExecuComp firms. Hartzell and Starks (2003) argue that concentrated ownership reflects the incentives of institutional owners to monitor executive compensation and that, as institutional ownership becomes more concentrated, institutions use their influence to increase the level of CEO incentives.

---

[2] We consider a firm to be predicted to be an ExecuComp firm if the probability generated by the model is greater than 50 percent.

[3] Note that a lower average weight on earnings in the non-ExecuComp sample does not mechanically lead to a smaller shift in the weight for the non-ExecuComp sample. To see this, consider the following example: Sample A contains a high-growth firm that places a 10 percent weight on earnings and a low-growth firm that places a 90 percent weight on earnings. Sample B contains a high-growth firm that places a 60 percent weight on earnings and a low-growth firm that places an 80 percent weight on earnings. In this example, the difference in the weights between high- and low-growth firms for Sample A is considerably larger than the difference in weights for Sample B even though the average weight on earnings for Sample A (50 percent) is lower than the average weight on earnings for Sample B (70 percent).

For non-ExecuComp firms, the lower degree of firm complexity and the greater noise in stock prices leads institutional shareholders to use alternative performance measures to evaluate management. We therefore predict that the positive relation between concentrated institutional ownership and CEO stock option grant incentives is stronger for ExecuComp firms than for non-ExecuComp firms.

Our results support this prediction. We find that the relation between institutional ownership concentration and the extent of CEO stock option grant incentives is significantly more positive for ExecuComp firms than it is for non-ExecuComp firms. We also find that the positive association between institutional ownership concentration and the extent of CEO stock option grant incentives is only present in the ExecuComp sample. Our findings are robust to adjustments for differences in CEO ownership and to using predicted classifications.

Finally, we examine whether the systematic differences in the functions that determine CEO pay that we document between ExecuComp and non-ExecuComp firms are also present within the ExecuComp sample by comparing S&P 500 firms to non-S&P 500 firms included in the ExecuComp database. We choose this separation because many of the differences in the contractual environment between ExecuComp and non-ExecuComp firms are also present between S&P 500 firms and non-S&P 500 ExecuComp firms. Given the cost of acquiring compensation data from other sources, these tests provide guidance to researchers regarding whether they can use ExecuComp data to assess the sensitivity of their results to differences in the contracting environment.

We find limited evidence that the differences we document across ExecuComp and non-ExecuComp firms can be found within the ExecuComp database. Specifically, with respect to our first prediction, we find that S&P 500 firms place a larger weight on stock returns, but we do not find a significant difference in the weight placed on accounting earnings by S&P 500 and non-S&P 500 firms within the ExecuComp sample. Also, we do not find significant differences between the two groups of firms with respect to the predicted negative association between the firm's investment opportunity set and the weight placed on earnings or the relation between institutional ownership concentration and CEO stock option grant incentives. In sum, these findings suggest that tests that compare groups of firms within the ExecuComp database should not be used as substitutes for tests that compare ExecuComp and non-ExecuComp firms to assess the impact of contracting environments on the design of CEO incentive contracts.

Overall, we document that ExecuComp firms are more complex, attract a greater degree of analyst following, have higher, but less concentrated institutional ownership, and have a greater degree of stock liquidity than non-ExecuComp firms. We refer to these factors collectively as the firm's contracting environment and examine whether the differences in contracting environments between ExecuComp and non-ExecuComp firms lead to predictable differences in the design of CEO incentive contracts. We find that ExecuComp firms place a greater weight on earnings and stock returns in setting CEO cash compensation, and that the weight placed on earnings is more sensitive to the extent of the firm's growth opportunities for ExecuComp firms than non-ExecuComp firms. Finally, the relation between institutional ownership concentration and CEO stock option grant incentives is significantly positive only for ExecuComp firms.

Our study contributes to the literature on the determinants of executive compensation contracts by providing insight into the design of CEO incentive contracts for a set of firms that have different contracting environment characteristics from those of firms typically examined in empirical studies on executive compensation. Our findings indicate that expanding samples to include

non-ExecuComp firms can substantially influence empirical results and provide researchers with additional insights into executive compensation and corporate governance issues.[4]

Finally, we find evidence to support the Baber et al. (1996) prediction regarding the impact of growth opportunities on the relative weight on earnings and the Hartzell and Starks (2003) prediction regarding the positive relation between institutional ownership concentration and CEO stock option grant incentives in the ExecuComp sample, but not in our non-ExecuComp sample. As a result, researchers should exercise caution when generalizing results based on an ExecuComp sample to broader samples.

Section II discusses our sample. Section III documents differences between ExecuComp and non-ExecuComp firms with regard to firm, CEO, and monitoring environment characteristics. Section IV develops empirical predictions based on these differences. Section V reports evidence on differences between ExecuComp and non-ExecuComp firm CEO compensation functions. Section VI provides the results of tests that compare S&P 500 firms with non-S&P 500 firms included in the ExecuComp database. Section VII concludes.

II. DESCRIPTION OF EXECUCOMP AND NON-EXECUCOMP FIRM SAMPLES

The Standard & Poor's ExecuComp database currently provides data from 1992 onward for the firms in the S&P 1500 index, which includes the S&P 500, S&P 400 mid-cap, and S&P 600 small-cap indices. From 2000 to 2007, ExecuComp firms comprise approximately 25 percent of the domestic firms in the Compustat database. Over this same period, the market capitalization of ExecuComp firms is approximately 88 percent of the market capitalization of all domestic Compustat firms.[5]

According to Standard & Poor's, the goal of the committees that determine the composition of the S&P 1500 indices is to ensure that the firms act as leading indicators of large-cap, medium-cap, and small-cap U.S. equities and have sufficient liquidity to ensure efficient portfolio trading. As a result, firms that are included must meet a number of criteria. One of these criteria is financial viability, which is usually determined by four consecutive quarters of positive GAAP earnings before discontinued operations and extraordinary items. S&P also requires each firm to have adequate stock liquidity. As a result of these criteria, the ExecuComp database is likely to include a relatively high proportion of large and profitable firms with stable cash flows and liquid shares.

We obtain compensation data for non-ExecuComp firms for the fiscal years 2000–2007 from the Morningstar Historical Governance database. This database can be purchased from Morningstar, and provides machine-readable compensation data for firms included in the Russell 3000 and S&P 1500 indices. Hence, the non-ExecuComp firms we study are included in the Russell 3000 index, but not in the S&P 1500 indices. The Russell 3000 index consists of firms in the Russell large-cap 1000 (the 1000 largest firms in the U.S. in terms of market capitalization in a given year), and the Russell small-cap 2000 (the next 2000 largest firms in the U.S. in terms of market capitalization in a given year). We note that according to Russell Investments, the Russell 2000 small-cap index makes up almost 10 percent of the market capitalization of U.S. equities. Thus,

---

[4] In addition to using data collected from the Morningstar Historical Governance database, our source for compensation data for non-ExecuComp firms, other data sources available to researchers wishing to study broader samples include Equilar and the Corporate Library. Further, in some instances hand-collection of compensation data for non-ExecuComp firms could be warranted and necessary to gain specific insights about compensation practices.

[5] Percentages were calculated using the sample of firms with data available on Compustat excluding firms incorporated abroad for the fiscal years 2000 through 2007. Market values are measured as the price of the firm's stock at the calendar year-end multiplied by the number of outstanding shares at calendar year-end.

our sample of non-ExecuComp firms enables us to study the compensation practices of a group of firms that we know little about, even though this group of firms is economically important.[6]

As shown in Panel A of Table 1, from 2000 through 2007, there are a total of 13,828 firm-years available on ExecuComp and a total of 17,480 firm-years available for Morningstar-non-ExecuComp firms. After deleting observations with insufficient data for the tests that follow, our final sample consists of 11,928 ExecuComp and 11,168 non-ExecuComp firm-years. Panel B provides information regarding the distribution of observations across the two samples by year. The number of Morningstar-non-ExecuComp firms increases over time. There is a large increase in this number of firms between 2001 and 2002. This large increase occurs because the years 2000 and 2001 are the first years for which the Morningstar Historical Governance database has comprehensive CEO compensation data and during these two years CEO compensation data was not collected for all the firms in the Russell 3000 and S&P 1500 indices. Subsequent to 2002, there is a gradual annual increase in the number of Morningstar-non-ExecuComp firms because once Morningstar covers a firm, they continue to track the firm as long as it is traded on a U.S. stock exchange. As a result, the sample of non-ExecuComp firms in the database increases over time, as firms are dropped from the Russell 3000 and new firms are added to this index. In contrast to the Morningstar-non-ExecuComp firms, the number of ExecuComp firms decreases over our sample period because, although ExecuComp likewise retains firms that are dropped from one of the S&P 1500 indices, it also backfills the data for new firms that are added to an index.

With respect to industry composition, Table 1, Panel C shows that the two sets of firms appear to be relatively similar, with the notable exceptions of a lower proportion of financial institutions (12.2 percent versus 25.8 percent) and a higher proportion of transportation and gas companies (10.9 percent versus 6.0 percent) and wholesale and retail firms (11.4 percent versus 6.9 percent) in the ExecuComp sample. The difference in the proportion of financial institutions is driven by the banking industry (two-digit SIC code 60), and the difference for transportation and gas companies is driven by the utility industry (two-digit SIC code 49).[7] The difference in the wholesale and retail sector is not driven by a specific two-digit SIC industry within this sector.

III. DIFFERENCES BETWEEN EXECUCOMP AND NON-EXECUCOMP FIRM CHARACTERISTICS

Table 2 compares firm, CEO, institutional ownership, and incentive compensation characteristics for the two samples, both for the full samples and after excluding banking and utility firms. In Panel A of Table 2 we report statistics of firm characteristics for firms in the ExecuComp, Morningstar-non-ExecuComp, and Compustat samples. Overall, non-ExecuComp firm characteristics significantly differ from those of ExecuComp firms. In addition, their characteristics more closely resemble those of the broad sample of Compustat firms than those of ExecuComp firms.

Given the selection criteria for inclusion in the three major S&P indices, it is not surprising to find, in Panel A of Table 2, that ExecuComp firms are substantially larger and more profitable than are Morningstar-non-ExecuComp firms. Median market capitalization (return on assets) for ExecuComp and non-ExecuComp firms is $1.56 billion dollars (4.5 percent) and $165 million dollars (1.1 percent), respectively. ExecuComp firms are also more highly leveraged and have lower

[6] When there was an overlap between ExecuComp and Morningstar data, we found agreement among the data sources for various data fields. In the analyses that follow, the ExecuComp sample is restricted to firms with data available from ExecuComp and the Morningstar non-ExecuComp sample is restricted to observations that do not appear in the ExecuComp sample.

[7] Approximately 5.4 percent of ExecuComp firms are in the banking industry, as opposed to 18.5 percent of firms in the Morningstar-non-ExecuComp sample. Also, about 6.0 percent of ExecuComp firms are utility firms, while only 1.8 percent of Morningstar-non-ExecuComp firms are utility firms.

TABLE 1

ExecuComp and Non-ExecuComp Samples

Panel A: Sample Selection Procedure (the sample period is 2000–2007)

	ExecuComp Firms	Morningstar Non-ExecuComp Firms	Total
Available firm-years	13,828	17,480	31,308
Missing necessary Compustat data items	(636)	(2,483)	(3,119)
Subtotal	13,192	14,997	28,189
Missing necessary CRSP data items	(638)	(3,272)	(3,910)
Subtotal	12,554	11,725	24,279
Missing institutional ownership data	(626)	(557)	(1,183)
Final sample	11,928	11,168	23,096

Panel B: Sample Observations by Year

Year	ExecuComp Firms	Morningstar Non-ExecuComp Firms	Total
2000	1,616	453	2,069
2001	1,543	825	2,368
2002	1,534	1,474	3,008
2003	1,574	1,563	3,137
2004	1,536	1,611	3,147
2005	1,492	1,721	3,213
2006	1,465	1,790	3,255
2007	1,168	1,731	2,899

(continued on next page)

Panel C: Percentage Industry Distribution of Samples

One-Digit SIC Code	Description	ExecuComp Firms	Morningstar Non-ExecuComp Firms	Compustat Firms
0	Agricultural	0.2	0.2	0.3
1	Metal, mining, and ore	4.9	4.3	4.5
2	Food, beverage, and chemical	16.3	14.8	15.4
3	Process and manufacturing	28.2	24.5	27.0
4	Transportation and gas	10.9	6.0	8.2
5	Wholesale and retail	11.4	6.9	9.7
6	Financial institutions	12.2	25.8	14.9
7	General services	12.1	12.6	15.2
8	Specialty services	3.5	4.8	4.6
9	Non-operating establishments	0.3	0.1	0.2
	Total	100.0	100.0	100.0

TABLE 2

Comparison of ExecuComp and Non-ExecuComp Firms Mean (Median)

Panel A: Firm Characteristics[a]

	Full Sample				Sample Excluding Banking and Utility Firms	
	ExecuComp	Non-ExecuComp	Combined Sample	Compustat	ExecuComp	Non-ExecuComp
Market value of equity ($000s)	7,817	637***	4,345	2,170	7,837	683***
	(1,558)	(165)***	(518)	(231)	(1,457)	(166)***
Book value of assets ($000s)	13,229	1,395***	7,507	2,900	10,318	1,311***
	(1,650)	(204)***	(624)	(335)	(1,298)	(149)***
Return on assets	3.4%	−6.1%***	−1.2%	−5.7%	3.5%	−7.3%***
	(4.5%)	(1.1%)***	(2.7%)	(1.4%)	(5.2%)	(1.3%)***
Return on equity	4.2%	−4.9%***	−0.0%	−6.9%	3.0%	−7.6%***
	(12.2%)	(7.2%)***	(9.9%)	(7.8%)	(12.0%)	(5.4%)***
Long-term debt/book value of assets	0.18	0.14***	0.16	0.16	0.18	0.14***
	(0.15)	(0.06)***	(0.10)	(0.09)	(0.15)	(0.05)***
Annualized stock return	12.6%	18.8%***	15.6%	10.0%	12.4%	21.1%***
	(6.9%)	(4.7%)***	(5.9%)	(2.8%)	(5.8%)	(5.4%)
Idiosyncratic risk	2.4%	3.3%***	2.8%	3.6%	2.5%	3.6%***
	(2.0%)	(2.8%)***	(2.4%)	(2.8%)	(2.1%)	(3.0%)***
Market-to-book equity	3.43	3.46***	3.45	2.68	3.60	3.75***
	(2.22)	(1.93)***	(2.08)	(1.85)	(2.30)	(2.02)***
IOS factor score	−0.013	−0.012***	−0.013	−0.014	−0.013	−0.012***
	(−0.015)	(−0.015)***	(−0.015)	(−0.016)	−0.015	−0.015***
Analyst coverage	9.5	2.8***	6.2	4.7	9.4	2.9***
	(8)	(1)***	(4)	(2)	(9)	(2)***
Number of reported business segments	5.7	3.9***	4.9	4.0	5.8	4.3***
	(4)	(3)***	(3)	(3)	(4)	(3)***
Number of reported geographic segments	6.5	4.2***	5.4	1.9	7.1	4.8***
	(6)	(3)***	(3)	(1)	(6)	(3)***
Herfindahl index of within-industry concentration	0.42	0.54***	0.43	0.50	0.39	0.47***
	(0.33)	(0.33)***	(0.33)	(0.33)	(0.33)	(0.33)***

(continued on next page)

Panel A: Firm Characteristics[a]

	Full Sample				Sample Excluding Banking and Utility Firms	
	ExecuComp	Non-ExecuComp	Combined Sample	Compustat	ExecuComp	Non-ExecuComp
Herfindahl index of geographic concentration	0.41 (0.31)	0.56*** (0.33)***	0.49 (0.33)	0.85 (1.00)	0.37 (0.28)	0.50*** (0.33)***
Illiquidity of firm's shares	0.076 (0.002)	2.590*** (0.092)***	1.291 (0.009)	2.948 (0.041)	0.085 (0.002)	2.589*** (0.084)

Panel B: CEO and Institutional Investor Ownership Characteristics[b]

	Full Sample			Sample Excluding Banking and Utility Firms	
	ExecuComp	Non-ExecuComp	Combined	ExecuComp	Non-ExecuComp
CEO age	55.5 (56)	54.5*** (54)***	55.2 (55)	55.4 (55)	54.3*** (54)***
CEO tenure	6.9 (5)	10.8*** (9)***	7.8 (5)	7.0 (5)	10.5*** (9)***
CEO ownership percentage	2.0% (0.2%)	5.7%*** (0.9%)***	3.8% (0.4%)	2.2% (0.3%)	6.2%*** (1.0%)***
Institutional ownership percentage	69.8% (73.2%)	36.3%*** (30.8%)***	53.6% (57.8%)	71.6% (75.5%)	38.6%*** (34.5%)***
Institutional ownership concentration	0.41 (0.38)	0.69*** (0.69)***	0.54 (0.47)	0.41 (0.39)	0.68*** (0.68)***

(continued on next page)

Panel C: CEO Incentive Compensation[b]

	Full Sample			Sample Excluding Banking and Utility Firms	
	ExecuComp	Non-ExecuComp	Combined	ExecuComp	Non-ExecuComp
Bonus as a percentage of total cash compensation	23.5% (21.0%)	18.3%*** (14.8%)***	20.9% (18.1%)	23.1% (20.2%)	18.3%*** (14.5%)***
Absolute percentage change in total cash compensation	36.8% (22.5%)	31.7%*** (15.7%)***	34.4% (18.9%)	37.4% (23.1%)	33.7%*** (17.0%)***
Value of annual option and restricted stock grants ($000s)	1,948 (657)	424*** (16)***	1,211 (171)	2,045 (707)	472*** (23)***
Ratio of value of annual option and restricted stock grants to total cash compensation	35.9% (36.2%)	19.7%*** (4.5%)***	28.1% (20.6%)	37.1% (38.3%)	21.2%*** (6.2%)***
Sensitivity of annual option and restricted stock grants for a 1% change in firm value ($000s)	58.8 (19.4)	17.1*** (0.6)***	38.6 (5.5)	60.7 (19.8)	19.0*** (0.7)***

*** Represents significance at the 1 percent level for two-tailed t-tests or Wilcoxon rank-sum tests to determine whether means or medians differ between the ExecuComp and non-ExecuComp samples.

[a] Data are taken from a sample of 11,928 observations from the ExecuComp database and 11,168 non-ExecuComp observations from the Morningstar Historical Governance database from 2000 through 2007. There are 37,362 observations in the Compustat sample for Compustat firms from 2000 through 2007 that are U.S.-listed firms that are not incorporated abroad and that have necessary data to calculate the firm characteristic variables that appear in Panel A. Market value of equity is the number of outstanding shares multiplied by the stock price at the end of the year. Book value of assets is the average book value of the firm's assets. Return on assets is net income before extraordinary items divided by average book assets. Return on equity is net income before extraordinary items divided by the average book value of equity. Long-term debt/book value of assets is as of the end of the fiscal year. Annualized stock return is the firm stock return based on the annualized monthly stock returns for the fiscal year, assuming dividend reinvestment. Idiosyncratic risk is the standard deviation of the residual from a market model regression estimated over 253 days using daily stock returns (subject to a minimum of 125 daily stock returns). Market-to-book equity is the market value of the firm's equity scaled by the book value of the firm's equity at the end of the fiscal year. IOS factor score represents the investment opportunity set based on investment intensity, market-to-book assets, geometric mean annual growth rate of market value of assets, and R&D expenditures scaled by assets following Baber et al. (1996). Analyst coverage data are collected from the Institutional Brokers' Estimate System (I/B/E/S). Herfindahl index of within-industry concentration is calculated by first determining what are the fractions of a firm's total sales accounted for by its different business segments and then summing the squares of these fractions. Herfindahl index of geographic concentration is calculated by first determining what are the fractions of a firm's total sales accounted for by its different geographic segments and then summing the squares of these fractions. Illiquidity of firm's shares is measured as in Amihud (2002) and represents the average daily ratio of stock return to dollar volume. Banking and utility firms are firms that operate in industries with SIC codes that begin with 60 or 49.

(continued on next page)

[b] CEO age, tenure, and ownership data are collected from the ExecuComp and Morningstar Historical Governance databases. Institutional ownership data are collected from the Thomson Financial database of 13f filings. Institutional ownership concentration is the fractional ownership held by the top five institutions scaled by the fraction of outstanding shares held by all institutions. Bonus is the annual cash bonus paid to the executive. Total cash compensation is salary plus cash bonus. Absolute percentage change in cash compensation is the absolute value of the change in total cash compensation scaled by total cash compensation in the prior fiscal year. Value of annual option grants is the Black-Scholes value of the options granted to the CEO throughout the fiscal year in thousands of dollars. Sensitivity of annual option and restricted stock grants for a 1 percent change in firm value ($000s) is the dollar change (in thousands of dollars) of the CEO's annual option and restricted stock grants for a 1 percent change in the value of the firm's equity.

idiosyncratic risk. Differences in growth prospects across the two samples are less clear. Non-ExecuComp firms have higher mean values for the market-to-book ratio and the investment opportunity set factor score (IOS), but ExecuComp firms have a higher median value for the market-to-book ratio.

Panel A of Table 2 also reports information on analyst coverage, firm complexity, and stock liquidity. Eight analysts follow the median ExecuComp firm, while only one analyst tracks the median non-ExecuComp firm. This suggests that there is likely to be greater public scrutiny of ExecuComp than non-ExecuComp firms. We also find that ExecuComp firms have more complex operations than non-ExecuComp firms, where complexity is measured as in Bushman et al. (2004), using a firm's reported number of business or geographic segments and Herfindahl-Hirschman indices measuring a firm's business or geographic diversification.[8] Further, using the Amihud (2002) measure of the illiquidity of a firm's stock (average daily ratio of stock return to dollar volume), we document that the shares of ExecuComp firms are significantly more liquid than those of non-ExecuComp firms. These differences persist after excluding banking and utility firms.

Table 2, Panel B presents a comparison of CEO and institutional ownership characteristics between the two samples. CEOs of ExecuComp firms are, on average, older and have shorter tenures. Also, CEOs of ExecuComp firms own a smaller percentage of the firm's outstanding shares. The median CEO ownership for ExecuComp and non-ExecuComp firms is 0.2 percent and 0.9 percent, respectively.

We also find significant differences in institutional ownership structure between the two groups of firms. Although institutions own approximately twice the proportion of outstanding shares of ExecuComp firms, this ownership is distributed across more institutions and is thereby less concentrated than the institutional ownership of non-ExecuComp firms. Because institutional monitoring should be more intense when institutional ownership concentration is higher (e.g., Shleifer and Vishny 1986; Hartzell and Starks 2003), institutions may monitor the management of non-ExecuComp firms with greater intensity than they would the management of ExecuComp firms.

Finally, Table 2, Panel C compares incentive compensation for CEOs across the ExecuComp and non-ExecuComp samples. We find that the bonus is a greater percentage of total cash compensation for firms in the ExecuComp sample, and that total compensation varies more from year to year. This panel also shows that ExecuComp firms grant a larger percentage of total compensation in the form of stock options and restricted stock. Further, this panel documents that the sensitivity of the value of the CEO's annual option and restricted stock grants to a 1 percent change in firm value is markedly higher for ExecuComp firms than it is for non-ExecuComp firms.

Table 3 provides evidence regarding the extent to which contracting environment characteristics explain the likelihood that a firm is included in the ExecuComp database. This analysis also addresses the concern that firm size fully captures the differences between ExecuComp and non-ExecuComp samples. Using the combined sample of ExecuComp and non-ExecuComp firms, we estimate a probit regression that predicts whether a firm is included in the ExecuComp database with variables that proxy for firm size, growth opportunities, the level and concentration of institutional ownership, analyst coverage, the number of reported business segments, and the illiquidity of a firm's stock. To separately evaluate the impact of firm characteristics correlated with firm size

[8] Because our sample period starts after 1998, when SFAS No. 131 was adopted, the number of business segments reported by a firm is unlikely to be affected by discretionary disclosure issues. Consequently, this number should be a valid measure of firm complexity.

TABLE 3
ExecuComp Prediction Model

Panel A: Probit Regression Model[a]

	(1)	(2)	(3)
Intercept	−0.091	−1.975***	−3.631***
	(−0.72)	(−8.89)	(−29.95)
Market-to-book equity	0.033***	0.029***	
	(4.11)	(3.51)	
IOS factor score	−17.237***	−18.666***	
	(−10.74)	(−11.54)	
Institutional ownership	1.193***	1.154***	
	(11.78)	(11.19)	
Institutional ownership concentration	−2.921***	−2.038***	
	(−19.02)	(−12.19)	
Analyst coverage	0.061***	0.024***	
	(12.60)	(4.65)	
Number of segments	0.061***	0.054***	
	(11.76)	(10.02)	
Illiquidity of firm's shares	0.000	0.003***	
	(0.25)	(2.90)	
Log of the market value of equity		0.268***	0.590***
		(10.14)	(30.59)
Pseudo R^2	0.422	0.440	0.335
n	23,096	23,096	23,096

Panel B: Prediction Errors Using p = 0.50 Cut-Off[b]

	Predicted Classification		
	ExecuComp	**Non-ExecuComp**	**Total**
Model Excluding Size			
True Classification			
ExecuComp	10,444	1,484	11,928
	(87.6%)	(12.4%)	
Non-ExecuComp	2,346	8,822	11,168
	(21.0%)	(79.0%)	
Total	12,790	10,306	23,096
Model Including Size			
True Classification			
ExecuComp	10,510	1,418	11,928
	(88.1%)	(11.9%)	
Non-ExecuComp	2,202	8,966	11,168
	(19.7%)	(80.3%)	
Total	12,712	10,384	23,096
Model Including Only Size			
True Classification			
ExecuComp	9,737	2,191	11,928
	(81.6%)	(18.4%)	

(continued on next page)

Panel B: Prediction Errors Using p = 0.50 Cut-Off[b]

	Predicted Classification		
	ExecuComp	**Non-ExecuComp**	**Total**
Non-ExecuComp	2,375 (11.0%)	8,793 (89.0%)	11,168
Total	12,112	10,984	23,096

*** Indicates that coefficients are different from zero at the 1 percent significance level.

[a] The sample period is 2000–2007. The models are probit regression models in which the dependent variable equals 1 if a firm is included in the ExecuComp database and equals 0 if a firm is a non-ExecuComp firm included in the Morningstar Historical Governance database. Market-to-book equity is the market value of the firm's equity at the end of the fiscal year scaled by the book value of the firm's equity. IOS factor score represents the investment opportunity set based on investment intensity, market-to-book assets, geometric mean annual growth rate of market value of assets, and R&D expenditures scaled by assets following Baber et al. (1996). Institutional ownership equals the number of shares held by institutions as a percentage of total outstanding shares. Institutional ownership concentration is the fractional ownership held by the top five institutions scaled by the fraction of outstanding shares held by all institutions. Logarithm of market value of equity is calculated at the end of the fiscal year. Illiquidity of firm's shares is measured as in Amihud (2002) and represents the average daily ratio of stock return to dollar volume. Year indicator variables are included in all the models. t-statistics for two-tailed tests are in parentheses and are based on standard errors that are adjusted for the clustering of observations at the firm level.

[b] A firm is predicted to be an ExecuComp firm if using the fitted values from the probit model for whether a firm is included in the ExecuComp database we find that the probability that the firm is an ExecuComp firm is at least 50 percent. In Panel B the rows provide information on the correct classification of ExecuComp and non-ExecuComp firms using the fitted values from the probit model. The columns provide information on the number of ExecuComp or non-ExecuComp firms predicted to be one or the other of the two firm types using the fitted values from the probit model.

on the classification of firms across the two samples, we first estimate the regression without size as an explanatory variable and then add size to the regression model. Finally, we run the regression using firm size as the only explanatory variable.

The results in the first column in Panel A of Table 3, in which size is excluded from the model, show that all of the independent variables except for the illiquidity of a firm's stock have a significant incremental effect in explaining whether a firm is included in the ExecuComp database. In the second column in Panel A of Table 3, we include the natural logarithm of the market value of equity as an explanatory variable. All of the significant associations documented in the first column are robust to the inclusion of firm size, indicating that differences in firm size between the two groups of firms do not completely drive the significant associations documented in the first column. Although the coefficient on firm size is positive and significant, as expected, the resulting increase in the explanatory power of the probit regression after including firm size, as measured by the pseudo R^2, is relatively small (0.422 to 0.440).

The third column in this panel shows that when firm size is the only explanatory variable in the probit model, the pseudo R^2 is 0.335. However, the increase in the pseudo R^2 when we add the other variables to the regression (0.335 to 0.440) further supports the contention that firm size is an important, but not all-encompassing variable in distinguishing ExecuComp and non-ExecuComp firms.

To further evaluate the economic determinants of the ExecuComp classification, we use the fitted values from the three models in Panel A of Table 3 to predict whether a firm is included in the ExecuComp database. We classify a firm as predicted to be an ExecuComp firm if the probability that the firm is an ExecuComp firm is greater than 50 percent. The results, shown in Panel B of Table 3, indicate that the first model in Panel A of Table 3 correctly classifies observations as

being ExecuComp observations 87.6 percent of the time. While this is substantially higher than a naïve prediction that every firm is an ExecuComp firm, which would be correct 51.6 percent (11,928/23,096) of the time given the distribution of our sample, some error still remains. Panel B also shows that the second model in Panel A of Table 3 correctly classifies ExecuComp firm observations 88.1 percent of the time. Thus, adding firm size to the prediction model slightly improves its predictive accuracy. Finally, the model that only includes firm size as an explanatory variable correctly classifies ExecuComp firm observations 81.6 percent of the time. Consistent with our interpretation of the results in Panel A of Table 3, the results presented here suggest that, although firm size is an important determinant of a firm's inclusion in the ExecuComp database, other characteristics are also important for predicting this likelihood.[9]

IV. EMPIRICAL PREDICTIONS

A compensation contract consists of a set of performance measures and weights that convert outcomes to pay, through mechanisms such as a bonus schedule. Theoretically, assuming costless contracting, an employment contract could contain an infinite number of performance measures, as long as each performance measure provides incremental information regarding the agent's actions (Holmstrom 1979). However, under costly contracting, a firm faces a cost-benefit trade-off when deciding whether to add a performance measure to the contract. Because the costs of adding performance measures are generally increasing in the number of existing performance measures, while the incremental benefit of adding performance measures decreases, firms typically use a relatively small number of performance measures in contract design.[10]

Firms commonly use aggregate financial measures, such as earnings or stock returns, because they are readily available, and are either audited (accounting earnings), or determined by independent parties (market returns). However, as noted by Ittner et al. (1997), firms can increase efficiency by supplementing the contract with nonfinancial performance measures, such as product quality measures or customer satisfaction (i.e., transferring a portion of the weight on the aggregate financial measures to the nonfinancial measures). The value of adding nonfinancial measures is determined by the firm's strategy, the incremental information provided by the nonfinancial measures, and the extent of exogenous noise in the financial measures.

In a relatively simple firm, the board can readily identify the firm's strategy and define performance measures that capture the key aspects to generating success, such as customer satisfaction. However, in a more complex firm, optimal strategies and keys to success are likely to differ across business units. For instance, while customer satisfaction may be the key to profitability in one business unit, increasing market share or minimizing product defect rates could be of greater importance in other business units. Thus, as the complexity of the firm increases, it becomes increasingly costly for a firm to balance conflicting strategies across business units and determine the appropriate weights to place on individual performance measures. Consequently, we expect that the more complex ExecuComp firms are more likely to rely on aggregate financial performance measures that allow their CEOs to use their presumably superior knowledge of firm operations to determine the optimal allocation of effort across activities and business units.

---

[9] We also use the combined sample of ExecuComp and non-ExecuComp firms to estimate the basic regressions found in Tables 4–6 on each size quartile. Overall, we do not find results across the size quartiles that are consistent with the results in these tables being driven by firm size.

[10] The cost of adding a performance measure includes the cost of determining what to measure, how to measure it, and the optimal weight to place on the measure. Because each of these steps requires a firm to assess the interaction of the proposed measure with all of the other measures in the contract, the cost of adding a performance measure increases with the number of measures in the contract.

At the same time, the higher analyst following and more liquid shares of ExecuComp firms reduce the exogenous noise in earnings and stock returns for these firms. Specifically, the additional scrutiny of ExecuComp firms by analysts leads these firms to increase their investment in information systems and reduces the ability of managers to manipulate earnings. Likewise, competition among analysts generates additional information that leads to the capital market having more precise information to evaluate the firm's earnings (Piotroski and Roulstone 2004). Further, the higher stock liquidity of ExecuComp firms improves the efficiency with which information is impounded into stock prices, and as a result, the stock returns of ExecuComp firms are likely to better reflect the value implications of the CEO's actions.

Overall, we expect the combination of the greater complexity of firm operations, higher analyst following, and greater stock liquidity of ExecuComp firms to increase the cost of using individual performance measures and reduce the noise associated with earnings and stock returns as measures of performance. As a result, we predict that ExecuComp firms exhibit stronger relations between CEO pay with earnings and stock returns than do non-ExecuComp firms.

The hypothesis that non-ExecuComp firms rely more heavily on individual performance measures and less on aggregate financial performance measures also has implications for shifts in the relative weights between earnings and stock returns. For example, Baber et al. (1996) and Ittner et al. (1997) argue that, due to the historical basis of earnings, as the importance of investment opportunities increases relative to assets in place, the value of accounting-based measures of performance declines. However, this reasoning assumes that firms with relatively low levels of investment opportunities place a relatively high weight on earnings. As noted above, non-ExecuComp firms are likely to substitute individual performance measures for earnings, even in the absence of growth opportunities. Thus, for non-ExecuComp firms, we would expect growth firms to shift the weights between individual performance measures rather than between aggregate financial performance measures. While the individual performance measures in CEO bonus contracts are generally not publicly disclosed, the underlying theory leads to the empirical prediction that the sensitivity of the weight on earnings to the firm's investment opportunity set should be greater for ExecuComp than non-ExecuComp firms.

Finally, we expect the differences in the contracting environment to affect the monitoring role played by institutional investors. Hartzell and Starks (2003) document a significantly positive relation between institutional ownership concentration and the extent of incentives provided to executives through annual stock option grants. Their interpretation is that institutional investors monitor the contracting process and use their influence to increase the degree of management incentives. They also argue that because institutions that monitor bear the full cost of the monitoring activity and share the benefit with the other shareholders, the extent of monitoring increases with the degree of concentration of institutional ownership, and an individual institution's influence over the board increases as its ownership becomes more concentrated. We expect the positive relation documented by Hartzell and Starks (2003) is less likely to hold for non-ExecuComp firms because the lower degree of complexity and the associated increase in the availability of individual performance measures for non-ExecuComp firms reduces the cost for institutions to directly evaluate management's decisions or use alternative performance measures to evaluate management's effectiveness. In addition, the expected higher degree of noise in stock prices for non-ExecuComp firms increases the noise in stock-based incentives, such as stock options, for those firms. Consequently, institutions that invest in non-ExecuComp firms are more likely to directly monitor management, or use individual performance measures to evaluate management. As a result, a positive relation between institutional ownership concentration and stock option grant incentives is less likely for these firms.

V. PRIMARY RESULTS

Weights on Aggregate Financial Performance Measures

To assess whether ExecuComp firms place greater weights on aggregate financial performance measures, we estimate the following regression relating changes in compensation to accounting and stock return measures of performance (see Lambert and Larcker 1987; Sloan 1993; Baber et al. 1996; Baber et al. 1999; Matsunaga and Park 2001).[11]

$$\Delta comp_{it} = \alpha_0 + \alpha_1 \Delta ROE_{it} + \alpha_2 \text{Annualized stock return}_{it} + \alpha_3 Exec + \alpha_4 (Exec * \Delta ROE_{it}) + \alpha_5 (Exec * \text{Annualized stock return}_{it}) + e_1 \quad (1)$$

where:

Δcomp = the annual change in the natural logarithm of the CEO's total cash compensation or total compensation;
ΔROE = the annual change in income before extraordinary items deflated by average stockholders equity;
Annualized stock return = annualized stock return assuming dividend reinvestment; and
Exec = 1 if the firm is in the ExecuComp database, and 0 otherwise.

We predict that ExecuComp firms place greater weights on aggregate financial performance measures in determining CEO compensation (i.e., α_4 and $\alpha_5 > 0$). Consistent with prior research, we exclude banking and utility firms because these firms are regulated and likely to compensate their CEOs differently from firms in unregulated industries. We first restrict the Exec indicator variable and interaction terms from the compensation model and estimate the model separately for ExecuComp and non-ExecuComp firms. We then expand the model and estimate Equation (1) on the pooled sample to test whether the relations between compensation and the performance metrics are significantly different between the samples.

Table 4, Panel A presents the results from estimating Equation (1), using CEO cash compensation as the dependent variable. We find evidence of the expected positive relation between CEO cash compensation and both accounting earnings and stock returns for both the ExecuComp sample (Column 1) and the non-ExecuComp sample (Column 2). For the pooled sample, the results reported in Column (3) indicate that the coefficients on the interactions of the ExecuComp indicator variable with the change in ROE and stock return variables are significantly positive. These findings suggest that ExecuComp firms place a greater weight on aggregate financial performance measures in determining cash compensation, which is consistent with our prediction that the greater complexity of ExecuComp firms increases the cost for these firms of using individual performance measures, while the higher analyst following and stock liquidity levels of ExecuComp firms reduce the noise in aggregate performance measures. At the same time, the relatively weaker relations between cash compensation and aggregate performance measures for the non-ExecuComp sample is consistent with firms excluded from the ExecuComp sample placing a greater weight on individual nonfinancial performance measures.

We conduct two additional tests to rule out alternative explanations. The first explanation is that there is a weaker relation between CEO cash compensation and aggregate financial perfor-

[11] See also the regression models in Dechow et al. (1994) and Gaver and Gaver (1998) for evidence on whether restructuring or other nonrecurring accounting charges modify the relation between cash compensation and earnings, Natarajan (1996) for results on how cash flows from operations modifies this relation, and Baber et al. (1998) for an examination of how earnings persistence affects this relation.

TABLE 4

Weights on Accounting and Stock Returns in Determining CEO Compensation

Panel A: Cash Compensation

	ExecuComp	Non-ExecuComp	Combined	Excluding Loss Firms	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−0.004	0.082***	−0.016	0.154***	−0.012	−0.023*
	(−0.27)	(5.01)	(−1.23)	(11.86)	(−0.93)	(−1.80)
ΔROE	0.232***	0.084***	0.080***	0.199***	0.083***	0.044**
	(9.40)	(3.70)	(3.52)	(5.67)	(3.51)	(2.22)
Annualized stock return	0.250***	0.127***	0.131***	0.149***	0.139***	0.121***
	(18.50)	(13.03)	(13.97)	(12.08)	(13.98)	(13.43)
ExecuComp indicator			−0.026***	−0.017**	−0.028***	−0.013**
			(−4.48)	(−2.32)	(−4.49)	(−2.13)
ExecuComp indicator * ΔROE			0.157***	0.149***	0.154***	0.266***
			(4.48)	(2.65)	(4.33)	(7.37)
ExecuComp indicator * Annualized stock return			0.113***	0.112***	0.106***	0.120***
			(6.99)	(5.66)	(6.41)	(7.81)
Adjusted R^2	0.044	0.053	0.083	0.099	0.086	0.090
n	7,342	5,935	13,277	9,852	12,597	13,277

Panel B: Total Compensation

	ExecuComp	Non-ExecuComp	Combined	Excluding Loss Firms	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−0.154***	−0.083*	−0.003	−0.023	−0.005	−0.023
	(−6.95)	(−1.90)	(−0.10)	(−0.93)	(−0.15)	(−0.77)
ΔROE	0.177***	0.056	0.049	1.46 **	0.028	0.103*
	(3.07)	(0.95)	(0.85)	(2.17)	(0.48)	(1.91)
Annualized stock return	0.500***	0.254***	0.256***	0.245***	0.286***	0.232***
	(12.48)	(10.49)	(11.25)	(8.12)	(13.05)	(10.23)

(continued on next page)

Panel B: Total Compensation

	ExecuComp	Non-ExecuComp	Combined	Excluding Loss Firms	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)	(6)
ExecuComp indicator			−0.080***	−0.077***	−0.081***	−0.061***
			(−5.57)	(−4.56)	(−5.42)	(−4.04)
ExecuComp indicator * ΔROE			0.132	0.052	0.151*	0.017
			(1.63)	(0.55)	(1.85)	(0.21)
ExecuComp indicator * Annualized stock return			0.238***	0.202***	0.211***	0.265***
			(5.13)	(4.36)	(4.55)	(6.25)
Adjusted R^2	0.072	0.031	0.051	0.053	0.057	0.051
n	7,342	5,935	13,277	9,852	12,597	13,277

*,**,*** Indicate that coefficients are different from zero at 10, 5, and 1 percent significance levels, respectively.

The sample period is 2000–2007. The dependent variable in the Panel A OLS regression models is the annual change in the logarithm of cash compensation (salary plus bonus) for each CEO. The dependent variable in the Panel B OLS regression models is the annual change in the logarithm of total compensation for each CEO. ROE is net income before extraordinary items divided by the average book value of equity. Annualized stock return is the firm stock return based on the annualized monthly stock returns for the fiscal year, assuming dividend reinvestment. ExecuComp indicator equals 1 if a firm is included in the ExecuComp database, and 0 otherwise. In Column (5) the observations for the highest CEO ownership decile from the non-ExecuComp sample are dropped from the analysis. In Column (6) the ExecuComp indicator variable is replaced with an indicator variable for whether a firm is predicted to be an ExecuComp firm with at least 50 percent probability using the fitted values from the second model in Table 3. Banking and utility firms have been omitted from the sample. All regressions include year fixed effects. t-statistics for two-tailed tests are in parentheses. Standard errors are adjusted for the clustering of observations at the firm level.

mance measures for loss firms and that the non-ExecuComp sample includes a higher proportion of loss firms (Gaver and Gaver 1998). To address this issue, we estimate the model on the combined sample after excluding loss firm-year observations. The results, reported in Column (4), show that excluding loss firms does not qualitatively affect our findings. The second alternative explanation is that the weaker relation between CEO cash compensation and aggregate financial performance measures is driven by differences between the two groups of firms with respect to CEO ownership. Because, as shown in Table 2, CEO ownership is greater for non-ExecuComp firms, the agency problem resulting from the separation of ownership and control could be lower for non-ExecuComp firms, thereby reducing the importance of providing the CEOs of these firms with contractual incentives. To address this concern, we estimate the model on the combined sample after excluding observations in the highest CEO ownership decile of the non-ExecuComp sample. This yields a sample for which there is not a significant difference in mean CEO ownership across ExecuComp and non-ExecuComp firms. The results, reported in Column (5), show that the significant and positive coefficients on the interactions of the ExecuComp indicator variable with the change in ROE and stock return variables are robust to the exclusion of non-ExecuComp firms with high CEO ownership.

Using the fitted values from the probit model in Column (2) of Table 3, Panel A, we investigate whether the larger weights on accounting earnings and stock returns for ExecuComp firms are due to differences in contracting environment characteristics between the two samples. Specifically, we replace the ExecuComp indicator variable with an indicator variable set equal to 1 if the probability that a firm is in ExecuComp based on its contracting environment characteristics is greater than 0.50, and 0 otherwise. The results, reported in Column (6) of Table 4, Panel A, are consistent with our other findings. We find that cash compensation is significantly more sensitive to earnings and stock returns for firms predicted to be in ExecuComp. These results suggest that the differences in the weights on the performance measures are due to contracting environments that are indicative of ExecuComp firms.

Table 4, Panel B reports the results using total compensation. Although adding equity grants provides a more comprehensive measure of CEO compensation, the extent to which equity grants are used to set long-term incentives (Core and Guay 1999), as opposed to providing a reward for short-term performance, will likely reduce the power of the test. As before, we find significant positive relations between stock returns and CEO compensation for both groups of firms, and also show that ExecuComp firms place a greater weight on stock returns in determining total compensation. We also find that, although the relation between accounting earnings and total CEO compensation is significantly positive for ExecuComp firms and not significant for non-ExecuComp firms, the coefficients on the change in ROE variable do not significantly differ between the two groups of firms. Columns (4) and (5) of Table 4, Panel B show that the results are not sensitive to the exclusion of loss firms or to the exclusion of high CEO ownership firms. Finally, the results reported in Column (6) show that the greater weight on stock returns for ExecuComp firms holds using the predicted classification.

The Impact of the Investment Opportunity Set on the Relative Weight on Earnings versus Returns

To test our prediction that the relative weight on earnings will be more sensitive to the investment opportunity set for ExecuComp firms, we estimate the following regression from Baber et al. (1996):

$$\Delta comp_{it} = \gamma_0 + \gamma_1 \Delta ROE_{it} + \gamma_2 \Delta ROE_{it} * IOS_{it} + \gamma_3 \text{Annualized stock return}_{it} + \gamma_4 \text{Annualized stock return}_{it} * IOS_{it} + \gamma_5 IOS_{it} + e_2 \quad (2)$$

where IOS is the factor score representing the investment opportunity set based on investment

intensity, market-to-book assets, geometric mean annual growth rate of market value of assets, and R&D expenditures scaled by assets following Baber et al. (1996).

As in Baber et al. (1996), banking and utility firms are excluded from the estimation. Baber et al. (1996) predict that the relative weight on earnings decreases as the investment opportunity set increases (i.e., $\gamma_2 < 0$).[12] We report the results for cash compensation and total compensation in Panels A and B of Table 5, respectively. Focusing on cash compensation, we find the predicted negative coefficient on the interaction of the investment opportunity set and accounting earnings for the ExecuComp sample in Column (1), but an insignificant coefficient on this interaction in the non-ExecuComp sample in Column (2). For the non-ExecuComp sample, we surprisingly find a significant negative coefficient on the interaction of the investment opportunity set and stock returns, which suggests that non-ExecuComp firms shift the weight away from stock returns as investment opportunities increase. In the pooled sample, reported in Column (3), we find that the coefficient on the ΔROE * IOS interaction is significantly more negative for the ExecuComp sample. However, the coefficient on the Annualized stock return * IOS interaction does not significantly differ between the two samples. Finally, the results reported in Columns (4)–(6) indicate that the inferences are robust to eliminating loss firms (Column (4)), controlling for differences in CEO ownership across samples (Column (5)), and using the predicted ExecuComp classification rather than the actual classification (Column (6)).[13]

For total compensation, reported in Table 5, Panel B, we find that for ExecuComp firms the coefficient on the ΔROE * IOS interaction is significant and negative and that the coefficient on the Annualized stock return * IOS interaction is significant and positive. Both of these findings are consistent with ExecuComp firms shifting away the weight on accounting earnings toward stock returns as the investment opportunity set increases. However, for non-ExecuComp firms, we find that the coefficients on these two interactions do not significantly differ. Finally, in the pooled sample (Column (3)), we find evidence that the coefficient on the Annualized stock return * IOS interaction is significantly more positive for ExecuComp firms than it is for non-ExecuComp firms. This difference remains significant when we adjust for differences in CEO ownership (Column (5)) or use the predicted classifications (Column (6)), although the difference is not significant when we exclude loss firms (Column (4)).

Overall, the Table 5 findings support the prediction that the sensitivity of the relative weight on earnings versus stock returns to the firm's investment opportunity set is greater for ExecuComp firms. Our tests also provide an example of how differences in the contracting environment can lead to inferences based on an ExecuComp sample that do not hold for a sample of non-ExecuComp firms.

The Determinants of Stock Option Grants

To test our prediction that the positive relation between institutional ownership concentration and CEO stock option grants will be stronger for ExecuComp firms we estimate the following model from Hartzell and Starks (2003):

---

[12] Technically, the prediction is that the weight on earnings relative to stock returns will be lower. This would be consistent with a significantly negative γ_2, and a positive, or insignificant γ_4, or a significantly positive γ_4 and a negative, or insignificant γ_2.

[13] As an alternative control for loss observations, in the Table 4 and Table 5 models we also tried including in the models run on only the ExecuComp sample or non-ExecuComp sample an indicator variable for whether earnings during the current year are negative and in the models run on the combined sample of ExecuComp and non-ExecuComp firms an indicator variable for whether earnings during the current year are negative and the interaction of this variable with the ExecuComp indicator variable. We found that including these control variables in these models did not affect the main findings from the models.

TABLE 5

The Impact of the Investment Opportunity Set on the Weights on Performance Measures

Panel A: Cash Compensation

	ExecuComp	Non-ExecuComp	Combined	Excluding Loss Firms	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−0.017	0.175**	0.047	−0.029	0.055	0.015
	(−0.52)	(2.36)	(1.33)	(−0.78)	(1.50)	(0.42)
ΔROE	1.193***	0.566***	0.540***	1.136***	0.561***	0.374***
	(9.71)	(4.56)	(4.37)	(5.49)	(4.37)	(3.23)
ΔROE * IOS	−0.363**	−0.087	−0.080	−0.192	−0.104	−0.055
	(−2.35)	(−0.81)	(−0.75)	(−0.83)	(−0.94)	(−0.05)
ExecuComp indicator * ΔROE			0.683***	0.767***	0.665***	1.129***
			(3.92)	(2.73)	(3.74)	(6.36)
ExecuComp indicator * ΔROE * IOS			−0.287**	−0.017*	−0.262**	−0.285**
			(−2.53)	(−1.65)	(−2.38)	(−2.49)
Annualized stock return	0.671***	0.382***	0.391***	0.451***	0.411***	0.308***
	(14.03)	(8.63)	(9.49)	(9.05)	(9.41)	(8.19)
Annualized stock return * IOS	0.007	−0.145***	−0.143***	−0.046	−0.156***	−0.129***
	(0.09)	(−3.49)	(−3.44)	(−0.54)	(−3.55)	(−3.13)
ExecuComp indicator * annualized stock return			0.268***	0.302***	0.251***	0.425***
			(4.47)	(4.23)	(4.08)	(7.35)
ExecuComp indicator * annualized stock return * IOS			0.147	0.222	1.61*	0.003
			(1.64)	(1.58)	(1.77)	(0.04)
IOS	−0.013	0.040**	0.041**	0.035	0.039*	0.033*
	(−0.49)	(1.97)	(2.06)	(0.97)	(1.82)	(1.65)
ExecuComp indicator * IOS			−0.057*	−0.065	−0.055	−0.007
			(−1.73)	(−1.28)	(−1.62)	(−0.19)
ExecuComp indicator			−0.023	0.009	−0.025	0.021
			(−1.09)	(0.36)	(−1.13)	(1.03)

(continued on next page)

Panel A: Cash Compensation

	ExecuComp	Non-ExecuComp	Combined	Excluding Loss Firms	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)	(6)
Adjusted R^2	0.078	0.031	0.057	0.072	0.059	0.064
n	7,342	5,935	13,277	9,852	12,608	13,277

Panel B: Total Compensation

	ExecuComp	Non-ExecuComp	Combined	Excluding Loss Firms	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−0.882***	−0.737	−0.415*	−0.243	−0.442*	−0.338
	(−3.51)	(−1.54)	(−1.78)	(−0.93)	(−1.84)	(−1.49)
ΔROE	1.026	0.175	0.107	0.305	−0.141	0.779
	(1.28)	(0.36)	(0.22)	(0.43)	(−0.27)	(1.54)
ΔROE * IOS	−2.585*	−0.020	−0.041	0.596	−0.162	−0.021
	(−1.67)	(−0.04)	(−0.09)	(0.40)	(−0.35)	(−0.05)
ExecuComp indicator * ΔROE			0.943	0.859	1.160	0.203
			(1.00)	(0.70)	(1.22)	(0.23)
ExecuComp indicator * ΔROE * IOS			−2.541	−1.072	−2.422	−1.366
			(−1.58)	(−0.40)	(−1.50)	(−1.11)
Annualized stock return	3.713***	1.524***	1.492***	1.469***	1.523***	1.099***
	(12.19)	(7.91)	(7.87)	(6.84)	(7.34)	(6.90)
Annualized stock return * IOS	2.411**	0.259	0.282	1.313***	0.309	0.093
	(3.20)	(1.31)	(1.42)	(2.84)	(1.45)	(0.50)
ExecuComp indicator * annualized stock return			2.244***	2.346***	2.216***	2.772***
			(6.36)	(5.96)	(6.12)	(8.71)
ExecuComp indicator * annualized stock return * IOS			2.073***	1.387	2.041***	1.536***
			(2.66)	(1.40)	(2.61)	(2.87)

(continued on next page)

Panel B: Total Compensation

	ExecuComp	Non-ExecuComp	Combined	Excluding Loss Firms	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)	(6)
IOS	−0.897***	−0.317***	−0.304***	−0.713***	−0.322***	−0.106
	(−4.17)	(−3.32)	(−3.19)	(−2.67)	(−3.24)	(−1.28)
ExecuComp indicator * IOS			−0.620***	−0.509	−0.598**	−0.999***
			(−2.63)	(−1.33)	(−2.52)	(−4.51)
ExecuComp indicator			−0.569***	−0.608***	−0.572***	−0.649***
			(−5.88)	(−5.23)	(−5.74)	(−6.76)
Adjusted R^2	0.063	0.028	0.050	0.052	0.052	0.054
n	7,342	5,935	13,277	9,852	12,608	13,277

*,**,*** Indicate that coefficients are different from zero at 10, 5, and 1 percent significance levels, respectively.

The sample period is 2000–2007. The dependent variable in the Panel A OLS regression models is the change in cash compensation (salary plus bonus) scaled by the prior year base salary for each CEO. The dependent variable in the Panel B OLS regression models is the change in total compensation scaled by the prior year total compensation for each CEO. Return on equity is net income before extraordinary items divided by the average book value of equity. Annualized stock return is the firm stock return based on the annualized monthly stock returns for the fiscal year, assuming dividend reinvestment. IOS is the investment opportunity factor score representing the investment opportunity set based on investment intensity, market-to-book assets, geometric mean annual growth rate of market value of assets, and R&D expenditures scaled by assets following Baber et al. (1996). ExecuComp indicator equals 1 if the firm is included in the ExecuComp database and equals 0 otherwise. In Column (5) the observations for the highest CEO ownership decile from the non-ExecuComp sample are dropped from the analysis. In Column (6) the ExecuComp indicator variable is replaced with an indicator variable for whether a firm is predicted to be an ExecuComp firm with at least 50 percent probability using the fitted values from the second model in Table 3. As in Baber et al. (1996), banking and utility firms have been omitted from the sample (observations with two-digit SIC codes 49 and 60). Year indicator variables are included in all the models. t-statistics for two-tailed tests are in parentheses and are based on standard errors that are adjusted for the clustering of observations at the firm level.

$$\begin{aligned}\text{Option}_{it} = {} & \theta_0 + \theta_1 \text{Institutional ownership concentration}_{it-1} + \theta_2 \text{Institutional ownership}_{it-1} \\ & + \theta_3 \text{Market-to-book assets}_{it-1} + \theta_4 \text{Log of the market value of equity}_{it-1} \\ & + \theta_5 \text{Change in shareholder wealth}_{it} + \theta_6 \text{Change in shareholder wealth}_{it-1} + e_3 \end{aligned} \quad (3)$$

where:

Option = the change in the value of the CEO's annual option grant for a \$1,000 increase in firm market value;
Institutional ownership concentration = the fractional ownership held by the top five institutions scaled by the fraction of outstanding shares held by all institutions;
Institutional ownership = the fraction of outstanding shares held by institutions;
Market-to-book assets = the market value of the firm's assets scaled by the book value of the firm;
Log of the market value of equity = the natural logarithm of the market value of the firm; and
Change in shareholder wealth = change in the value of the shares outstanding times stock price from period $t-1$ to period t.

Columns (1) and (2) of Table 6 provide the results when estimating Equation (3) on the ExecuComp and non-ExecuComp samples, respectively. In Column (1), consistent with Hartzell and Starks (2003), we find a significantly positive coefficient on institutional ownership concentration. However, Column (2) shows that this association is negative and significant for non-ExecuComp firms, which suggests that for these firms more intense institutional monitoring leads to a reduced flow of equity incentives to the CEO. The results for the pooled sample (Column (3)), where we include interactions of each independent variable with the ExecuComp firm indicator variable, show that the coefficient on institutional ownership concentration is significantly more positive for ExecuComp firms. This finding holds after controlling for differences in CEO ownership across samples (Column (4)) and using predicted classifications (Column (5)). These results are consistent with our prediction that for non-ExecuComp firms the lower complexity of operations combined with the availability of alternative performance measures reduces the positive association between institutional ownership concentration and CEO stock option grant incentives.

VI. TESTS WITHIN THE EXECUCOMP SAMPLE

The prior series of tests document significant differences in the structure of CEO compensation between ExecuComp and non-ExecuComp firms. We next examine whether similar differences can be found between broad groups of firms within the ExecuComp data set. Specifically, we compare the CEO compensation functions of S&P 500 firms with those of non-S&P 500 firms included in the ExecuComp sample. These tests provide guidance to researchers regarding whether ExecuComp data can be used to assess the sensitivity of results to different contracting environments.

Table 7 provides descriptive firm, monitoring, and compensation statistics for the three indices within the S&P 1500. The statistics suggest similar systematic differences in the contracting environment and compensation characteristics across firms in the ExecuComp sample. Firm size, firm complexity, analyst coverage, stock liquidity, the IOS factor score, and the market-to-book ratio all decrease as one moves from the S&P 500 index to the S&P 400 mid-cap and the S&P 600 small-cap indices. Further, CEO ownership and institutional ownership concentration increases

TABLE 6

The Pay-for-Performance Sensitivity of Option Grants

	ExecuComp	Non-ExecuComp	Combined	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)
Intercept	4.647***	8.583	12.046***	8.836***	13.991***
	(3.47)	(1.19)	(2.74)	(3.60)	(2.64)
Log of the market value of equity	−0.511***	−1.616***	−1.599***	−1.029***	−1.869***
	(−12.57)	(−2.68)	(−3.03)	(−4.07)	(−2.93)
Market-to-book assets	−0.002	0.035	0.003	0.007	0.003
	(−0.12)	(0.78)	(0.11)	(0.38)	(0.11)
Change in shareholder wealth	0.028	−0.092	−0.058	−0.044	−0.087
	(0.58)	(−0.19)	(−1.01)	(−0.90)	(−1.49)
Prior year change in shareholder wealth	0.032	0.531	0.062	0.046	−0.007
	(0.70)	(1.14)	(1.06)	(1.02)	(−0.13)
Institutional ownership concentration	0.827*	−9.958***	−7.746***	−5.221***	−8.347**
	(1.70)	(−2.66)	(−2.65)	(−3.80)	(−2.46)
Institutional ownership	1.652***	5.385***	3.030***	1.525**	3.357***
	(5.46)	(3.74)	(3.66)	(2.36)	(3.75)
ExecuComp indicator			−11.816**	−7.471***	−13.113**
			(−2.40)	(−3.30)	(−2.28)
ExecuComp indicator * Log of the market value of equity			1.312**	0.670***	1.629**
			(2.26)	(2.59)	(2.35)
ExecuComp indicator * Market-to-book assets			−0.000	−0.000	−0.000
			(−0.64)	(−1.01)	(−0.45)
ExecuComp indicator * Change in shareholder wealth			0.000	0.000	0.000
			(0.74)	(0.88)	(1.38)
ExecuComp indicator * Prior year change in shareholder wealth			−0.000**	0.000**	−0.000*
			(−1.98)	(−1.99)	(−1.91)

(continued on next page)

TABLE 6 (continued)

	ExecuComp	Non-ExecuComp	Combined	Excluding High CEO Ownership	Predicted Classifications
	(1)	(2)	(3)	(4)	(5)
ExecuComp indicator * Institutional ownership concentration			7.608***	5.400***	6.439**
			(2.71)	(3.40)	(2.06)
ExecuComp indicator * Institutional Ownership			0.989	1.507*	0.563
			(1.08)	(1.95)	(0.49)
Adjusted R^2	0.011	0.013	0.008	0.009	0.008
n	11,928	11,168	23,096	21,980	23,096

*,**,*** Indicate that coefficients are different from 0 at the 10, 5, and 1 percent significance levels, respectively.

The sample period is 2000–2007. The dependent variable in each Tobit regression model is the change in the value of the annual option grant to the CEO per $1,000 change in shareholder wealth. Logarithm of market value of equity is the prior year value and is calculated at the end of the fiscal year. Market-to-book assets is the prior year market value of the firm's assets scaled by the book value of the firm at the end of the fiscal year. Change in shareholder wealth is the change in value of the shares outstanding times stock price from period $t-1$ to period t. Institutional ownership concentration is the prior year fractional ownership held by the top five institutions scaled by the prior year fraction of outstanding shares held by all institutions. Institutional ownership is the prior year fraction of outstanding shares held by institutions. In Column (4) the observations for the highest CEO ownership decile from the non-ExecuComp sample are dropped from the analysis. In Column (5) the ExecuComp indicator variable is replaced with an indicator variable for whether a firm is predicted to be an ExecuComp firm with at least 50 percent probability using the fitted values from the second model in Table 3. ExecuComp indicator equals 1 if a firm is included in the ExecuComp database and equals 0 otherwise. Two-digit SIC code industry and year indicator variables are included in all the models. t-statistics for two-tailed tests are in parentheses and are adjusted for the clustering of observations at the firm level.

TABLE 7

Comparison of S&P 500 Large-Cap, S&P 400 Mid-Cap, and S&P 600 Small-Cap Firms Mean (Median)

Panel A: Firm Characteristics[a]

	S&P 500 Large-Cap	S&P 400 Mid-Cap	S&P 600 Small-Cap
Market value of equity	22,865	2,153	764***
	(9,045)	(1,889)	(621)***
Book value of assets	34,230	3,180	1,112***
	(8,611)	(1,779)	(579)***
Return on assets	6.3%	5.9%	4.8%***
	(5.9%)	(5.4%)	(4.9%)***
Return on equity	17.1%	13.4%	10.3%***
	(16.3%)	(13.3%)	(11.1%)***
Long-term debt/book value of assets	0.19	0.18	0.15***
	(0.17)	(0.17)	(0.11)***
Annualized stock return	13.7%	14.4%	14.8%***
	(8.9%)	(9.4%)	(7.0%)***
Idiosyncratic risk	1.9%	2.0%	2.5%***
	(1.6%)	(1.8%)	2.2%***
Market-to-book equity	3.88	2.97	2.51***
	(2.99)	(2.33)	(2.01)***
IOS Factor Score	−0.011	−0.012	−0.013***
	(−0.013)	(−0.014)	(−0.016)***
Analyst coverage	15.6	8.4	5.4***
	(15)	(7)	(5)***
Number of reported business segments	6.3	5.8	4.9***
	(5)	(5)	(3)***
Number of reported geographic segments	7.8	6.3	6.3***
	(6)	(6)	(4.5)***
Herfindahl index of within-industry concentration	0.42	0.40	0.41***
	(0.33)	(0.32)	(0.33)***
Herfindahl index of geographic concentration	0.36	0.40	0.41***
	(0.25)	(0.32)	(0.31)***
Illiquidity of firm's shares	0.001	0.004	0.019***
	(0.000)	(0.001)	(0.005)***

Panel B: CEO and Institutional Investor Ownership Characteristics[b]

	S&P 500 Large-Cap	S&P 400 Mid-Cap	S&P 600 Small-Cap
CEO age	55.9	56.0	55.9***
	(56)	(56)	(56)***
CEO tenure	6.1	7.6	8.0***
	(4)	(5)	(5)***
CEO ownership percentage	0.8%	2.7%	2.6%***
	(0.1%)	(0.3%)	(0.4%)***
Institutional ownership percentage	69.4%	72.5%	72.3%***
	(70.0%)	(75.7%)	(76.4%)***
Institutional ownership concentration	0.35	0.38	0.42***
	(0.34)	(0.37)	(0.40)***

(continued on next page)

Panel C: CEO Incentive Compensation[b]

	S&P 500 Large-Cap	S&P 400 Mid-Cap	S&P 600 Small-Cap
Bonus as a percentage of total cash compensation	29.0% (26.0%)	28.0% (25.9%)	25.6%*** (23.5%)***
Absolute percentage change in total cash compensation	36.5% (19.4%)	34.6% (20.3%)	30.3%*** (18.7%)***
Value of annual option and restricted stock grants ($000s)	3,678.1 (2,112.1)	1,371.9 (678.9)	699.6*** (252.3)***
Ratio of value of annual option and restricted stock grants to total cash compensation	45.6% (49.0%)	34.6% (34.8%)	28.0%*** (23.6%)***
Sensitivity of annual option and restricted stock grants for a 1% change in firm value ($000s)	109.7 (65.8)	42.4 (20.9)	20.2*** (7.3)***

[a] *** Represents significance at the 1 percent level for two-tailed t-tests or Wilcoxon rank-sum tests to determine whether means or medians differ between the S&P 500 large-cap and S&P 600 small-cap samples.

Data are taken from a sample of 3,209 S&P 500 large-cap firm-year observations, 2,204 S&P 400 mid-cap firm-year observations and 2,738 S&P 600 small-cap firm-year observations from 2000–2007. All firms are included on the ExecuComp database. Market value of equity is the number of outstanding shares multiplied by the stock price at the end of the year. Book value of assets is the average book value of the firm's assets. Return on assets is net income before extraordinary items divided by average book assets. Return on equity is net income before extraordinary items divided by the average book value of equity. Long-term debt/book value of assets is as of the end of the fiscal year. Annualized stock return is the firm stock return based on the annualized monthly stock returns for the fiscal year, assuming dividend reinvestment. Idiosyncratic risk is the standard deviation of the residual from a market model regression estimated over 253 days using daily stock returns (subject to a minimum of 125 daily stock returns). Market-to-book equity is the market value of the firm's equity scaled by the book value of the firm's equity at the end of the fiscal year. IOS factor score represents the investment opportunity set based on investment intensity, market-to-book assets, geometric mean annual growth rate of market value of assets, and R&D expenditures scaled by assets following Baber et al. (1996). Analyst coverage data is collected from the Institutional Brokers' Estimate System (I/B/E/S). Herfindahl index of within-industry concentration is calculated by first determining what are the fractions of a firm's total sales accounted for by its different business segments and then summing the squares of these fractions. Herfindahl index of geographic concentration is calculated by first determining what are the fractions of a firm's total sales accounted for by its different geographic segments and then summing the squares of these fractions. Illiquidity of firm's shares is measured as in Amihud (2002) and represents the average daily ratio of stock return to dollar volume.

[b] *** Represents significance at the 1 percent level for two-tailed t-tests or Wilcoxon rank-sum tests to determine whether means or medians differ between the ExecuComp and non-ExecuComp samples. Institutional ownership concentration is the fractional ownership held by the top five institutions scaled by the fraction of outstanding shares held by all institutions. Bonus is the annual cash bonus paid to the executive. Total cash compensation is salary plus cash bonus. Absolute percentage change in cash compensation is the absolute value of the change in total cash compensation scaled by total cash compensation in the prior fiscal year. Value of annual option grants is the Black-Scholes value of the options granted to the CEO throughout the fiscal year in thousands of dollars. Sensitivity of annual option and restricted stock grants for a 1 percent change in firm value ($000s) is the dollar change (in thousands of dollars) of the CEO's annual option and restricted stock grants for a 1 percent change in the value of the firm's equity.

across the three groups of firms, while the use of incentive-based compensation decreases. However, although there appears to be heterogeneity within the ExecuComp sample, the small-cap firms in the ExecuComp database more closely resemble ExecuComp firms than non-ExecuComp firms.

Table 8 compares the results from the ExecuComp/non-ExecuComp classification to the S&P 500/non-S&P 500 classification within the ExecuComp database. These tests provide evidence as to whether the differences in compensation functions between ExecuComp and non-ExecuComp

TABLE 8

S&P 500 Indicator within the ExecuComp Sample

Panel A: Performance Measure Weights for Cash Compensation (Table 4 Panel A)

	ExecuComp Indicator	S&P 500 Indicator
Indicator variable * ΔROE	0.157**	0.129
	(4.48)	(1.59)
Indicator variable * Annualized stock return	0.113***	0.098***
	(6.99)	(2.68)
n	13,277	7,342

Panel B: The Impact of the IOS on the Weight on Earnings in Cash Compensation (Table 5 Panel A)

	ExecuComp Indicator	S&P 500 Indicator
Indicator variable * ΔROE	0.683***	0.490
	(3.92)	(1.25)
Indicator variable * ΔROE * IOS	−0.287**	0.186
	(−2.53)	(0.37)
Indicator variable * Annualized stock return	0.268***	0.392**
	(4.47)	(2.98)
Indicator variable * Annualized stock return * IOS	0.147	0.181
	(1.64)	(0.69)
Indicator variable * IOS	−0.057*	−0.027
	(−1.73)	(−0.38)
n	13,277	7,342

Panel C: The Pay-For-Performance Sensitivity of Option Grants (Table 6)

	ExecuComp Indicator	S&P 500 Indicator
Indicator variable * Log of the market value of value of equity	1.312**	0.315***
	(2.26)	(3.03)
Indicator variable * Market-to-book assets	−0.000	−0.005
	(−0.64)	(−0.66)
Indicator variable * Change in shareholder wealth	0.000	−0.000
	(0.74)	(−0.11)
Indicator variable * Prior year change in shareholder wealth	−0.000**	−0.000***
	(−1.98)	(−2.89)
Indicator variable * Institutional ownership concentration	7.608***	−2.131
	(2.71)	(−1.59)
Indicator variable * Institutional Ownership	0.989	−0.840
	(1.08)	(−1.41)
n	23,096	11,928

,* Indicate that coefficients are different from zero at the 5 and 1 percent significance levels, respectively.

Results in the ExecuComp Indicator column provide regression coefficients for selected variables from the models in Table 4, Panel A, Table 5, Panel A, and Table 6. Results in the S&P 500 Indicator column provide regression coefficients for selected variables from the models in Table 4, Panel A, Table 5, Panel A, and Table 6 in which only the ExecuComp sample is examined and the ExecuComp indicator variable is replaced with an S&P 500 indicator variable. t-statistics for two-tailed tests are in parentheses and are based on standard errors that are adjusted for the clustering of observations at the firm level.

firms are present within the ExecuComp database. For brevity, we only tabulate the results for the key interactions of interest. To facilitate comparison with our prior findings we also include the comparable results from Tables 4–6.

In Table 8, Panel A, consistent with our results comparing ExecuComp and non-ExecuComp firms, we find that cash compensation is more sensitive to stock returns for S&P 500 than for non-S&P 500 firms. On the other hand, we do not observe the significant difference in the weight on earnings that we found between ExecuComp and non-ExecuComp firms. Table 8, Panel B shows that we do not find significant differences between S&P 500 and non-S&P 500 firms with respect to the predicted negative association between the firm's investment opportunity set and the weight placed on earnings versus stock returns. Likewise, Table 8, Panel C documents that the relation between institutional ownership concentration and CEO stock option grant incentives is not significantly different across the two groups of firms.[14] Overall, these findings indicate that tests conducted between groups of firms within the ExecuComp database are not substitutes for tests conducted between samples of ExecuComp and non-ExecuComp firms in assessing the impact of contracting environments on the design of CEO incentive contracts.

VII. CONCLUSION

The ExecuComp database provides an easy-to-use data source of a relatively broad range of firms, including the largest and arguably most important firms in the economy. As a result, the database is extensively used to investigate a wide range of governance and compensation issues, although little is known with regard to the governance structures and compensation functions of firms not included in the database, or how they differ from those of firms included in ExecuComp. In this study we document systematic differences in contracting environment characteristics between ExecuComp and non-ExecuComp firms that are likely to impact firms' executive compensation contracts. We hypothesize that the greater firm complexity, analyst following, and stock liquidity associated with ExecuComp firms alters the contracting environment for these firms, which leads to three empirical predictions. First, we predict that ExecuComp firms place a greater weight on earnings and stock returns in CEO cash compensation contracts. Second, we predict that the weight on earnings will be more sensitive to a firm's growth opportunities for ExecuComp firms. Third, we predict that the positive association between institutional ownership concentration and CEO stock option incentives documented by Hartzell and Starks (2003) will be stronger for ExecuComp firms than for non-ExecuComp firms.

The findings support our hypotheses and suggest that differences in contracting environments between ExecuComp and non-ExecuComp firms lead to differences in the structure of CEO incentive contracts and the role played by external monitors, such as institutional shareholders, in setting CEO incentives. Overall, our findings highlight the importance for future researchers to consider how a firm's contracting environment can affect results of empirical tests of executive compensation. Researchers should use caution when generalizing results to firms that are not included in the ExecuComp data set.

REFERENCES

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5: 31–56.

[14] For the Table 8 analyses we define whether a firm is an S&P 500 firm by considering whether it is included in the S&P 500 index during the most recent year for which ExecuComp has collected executive compensation data. If instead, for these analyses, we use an indicator variable that indicates whether in a particular year a firm is included in this index, we find similar results.

Baber, W. R., S. N. Janakiraman, and S. Kang. 1996. Investment opportunities and the structure of executive compensation. *Journal of Accounting and Economics* 21: 297–318.

———, S. Kang, and K. R. Kumar. 1998. Accounting earnings and executive compensation: The role of earnings persistence. *Journal of Accounting and Economics* 25: 169–193.

———, ———, and ———. 1999. The explanatory power of earnings levels vs. earnings changes in the context of executive compensation. *The Accounting Review* 74: 459–472.

Banker, R., and S. Datar. 1989. Sensitivity, precision, and linear aggregation of accounting signals. *Journal of Accounting Research* 27: 21–39.

Bushman, R., Q. Chen, E. Engel, and A. Smith. 2004. Financial accounting information, organizational complexity and corporate governance systems. *Journal of Accounting and Economics* 37: 167–201.

Core, J. E., and W. R. Guay. 1999. The use of equity grants to manage optimal equity incentive levels. *Journal of Accounting and Economics* 28: 151–184.

Dechow, P. M., M. R. Huson, and R. G. Sloan. 1994. The effect of restructuring charges on executives' cash compensation. *The Accounting Review* 69: 138–156.

Gaver, J. J., and K. M. Gaver. 1998. The relation between nonrecurring accounting transactions and CEO cash compensation. *The Accounting Review* 73: 235–253.

Hartzell, J. C., and L. T. Starks. 2003. Institutional investors and executive compensation. *The Journal of Finance* 58: 2351–2374.

Holmstrom, B. 1979. Moral hazard and observability. *The Bell Journal of Economics* 10: 74–91.

Ittner, C. D., D. F. Larcker, and M. V. Rajan. 1997. The choice of performance measures in annual bonus contracts. *The Accounting Review* 72: 231–255.

Lambert, R. A., and D. F. Larcker. 1987. An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research* 25: 201–227.

Matsunaga, S. R., and C. W. Park. 2001. The effect of missing a quarterly earnings benchmark on the CEO's annual bonus. *The Accounting Review* 26: 85–125.

Natarajan, R. 1996. Stewardship value of earnings components: Additional evidence on the determinants of executive compensation. *The Accounting Review* 71: 1–22.

Piotroski, J., and D. Roulstone. 2004. The influence of analysts, institutional investors, and insiders on the incorporation of market, industry and firm-specific information into stock prices. *The Accounting Review* 79: 1119–1152.

Shleifer, A., and R. Vishny. 1986. Large shareholders and corporate control. *The Journal of Political Economy* 94: 461–488.

Sloan, R. G. 1993. Accounting earnings and top executive compensation. *Journal of Accounting and Economics* 16: 55–100.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1545–1575

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1545

Ranking Performance Measures in Multi-Task Agencies

Peter O. Christensen
Aarhus University

Florin Şabac
Jie (Joyce) Tian
University of Alberta

ABSTRACT: We derive sufficient conditions for ranking performance evaluation systems in multi-task agency models (using both optimal and linear contracts) in terms of a second-order stochastic dominance (SSD) condition on the likelihood ratios. The SSD condition can be replaced by a variance-covariance matrix of likelihood ratios (VCM) condition when the utility function is square-root, the performance measures are normally distributed, and for LEN models. We identify existing results derived under the LEN assumptions that rely on the VCM condition and, thus, also hold for optimal contracts.

Keywords: *likelihood ratios; multiple tasks; LEN.*

I. INTRODUCTION

The choice of performance evaluation systems is at the heart of accounting and accounting research. The principal-agent framework has provided profound insights into the determinants of the optimal choice among alternative systems and how the information is optimally exploited in the associated reward system, such as the identification of the *conditional* controllability principle for the value of an additional performance measure (e.g., Antle and Demski 1988). However, most of the research in this area either is in single-period and single-task settings, or has extended the analysis to multiperiod and multi-task settings using the simplifying assumptions of the LEN framework (Linear contracts, Exponential utility, Normally distributed performance measures). The multi-task models of Feltham and Xie (1994) and Holmström and Milgrom (1991) are prominent examples of LEN models that have had a significant impact on

We thank the associate editor and two anonymous reviewers for their comments and suggestions. We also thank seminar participants at the Minnesota Theory Conference 2007, the EIASM Workshop on Accounting and Economics 2008, and at the University of British Columbia for their comments. Professor Şabac gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: May 2008
Accepted: December 2009
Published Online: August 2010

accounting research, broadening our understanding of performance measure design and other organizational issues.[1]

The advances made by using LEN models came at a price: (1) basic questions on the choice of performance evaluation systems remained unanswered in general; and (2) the validity of the results obtained with linear contracts is uncertain because performance information is not exploited optimally in the reward system. Our study speaks to both issues. We investigate the choice among performance evaluation systems in a single-period, multi-task setting with optimal contracts. The likelihood ratios reveal the underlying information economics not only for optimal contracts, but also for the LEN framework. We do not claim that LEN results have general validity, but we identify a set of LEN results for which the assumption of linear contracts is not restrictive.

We determine sufficient conditions for a partial ranking of information systems in multi-task agencies in terms of a second-order stochastic dominance (SSD) condition on the likelihood ratios. In general, the SSD condition is implied by Blackwell garbling. For normally distributed performance measures, the SSD condition is also implied by a dominance condition for the variance-covariance matrices (VCM) of the likelihood ratios. The same VCM condition is sufficient to rank information systems when the agent has square-root utility or in the LEN framework. Using the VCM condition, we extend Holmström's (1979) sufficient statistic result to the multi-task setting of normally distributed performance measures with both optimal and LEN contracts.

We provide the first results in the literature on ranking information systems for multi-task agency problems using optimal contracts. Our SSD ranking condition for the multi-task setting generalizes that in Kim (1995) for single-task settings. The sufficiency of the VCM dominance condition for multi-task settings and either square-root agent utility or normally distributed performance measures generalizes the single-task results in Kim and Suh (1991). The sufficiency of the VCM condition for ranking information systems in the LEN model is a new result based on a novel derivation of the optimal linear contracts using likelihood ratios.

In the single-task setting of Kim (1995), the SSD condition on the likelihood ratio reflects that a larger variation in the likelihood ratio provides more precise inferences about the agent's action and, thus, the risk premium required to induce a particular action can be reduced. In the extreme, if there is no variation in the likelihood ratio, the performance measure provides no information about the agent's action and, consequently, the performance measure cannot be used to provide incentives for the agent.

The same interpretation applies in our multi-task setting, except that we require a SSD condition on the (multi-dimensional) likelihood ratio weighted by the optimal multiplier on the incentive compatibility constraints. Because these multipliers are endogenously determined, our general SSD condition depends on the specifics of the incentive problem, such as the agent's preferences. This is the key difference relative to the single-task setting.

In the single-task setting of Kim (1995), the SSD condition on the (one-dimensional) likelihood ratios is independent of the multipliers because a stochastic dominance relation between distributions of two random variables is preserved if both variables are scaled by the same factor. In other words, the multiplier "cancels out," and the ranking of performance measures is based

[1] Despite the fact that linear contracts are not optimal in the many cases of interest to which the LEN model has been applied, its tractability has allowed exploration of many interesting contracting settings that would have been otherwise intractable. A partial list includes: Gibbons and Murphy (1992) on career concerns; Meyer and Vickers (1997) on relative performance evaluation; Bushman and Indjejikian (1993), Kim and Suh (1993), and Feltham and Wu (2000) on earnings and stock prices as performance measures; Datar et al. (2001) on relative incentive rates; Dutta and Reichelstein (1999, 2002, 2003) on dynamic performance measurement; Indjejikian and Nanda (1999), Christensen et al. (2003, 2005), and Şabac (2007, 2008) on dynamic models with renegotiation; McLaughlin (1994), Huddart and Liang (2005), Adams (2006), and Liang et al. (2008) on team incentives.

only on the informational characteristics of the performance measures. For general multi-task agencies, the optimal multipliers enter the ranking condition for performance evaluation systems.[2]

We provide several sufficient conditions such that the ranking of performance evaluation systems depends only on the informational characteristics of the performance measures. First, we show that in a setting with stochastically independent task-specific (technologically independent) performance measures, a SSD condition for each of the task-specific likelihood ratios implies the general SSD condition (independently of the optimal multipliers). In other words, if there is more variation in the likelihood ratio for each task, then the risk premium required for inducing a particular multi-task action can be reduced.

Second, we extend Kim (1995) and show that Blackwell garbling implies the general SSD condition for the weighted likelihood ratios in multi-task settings independently of the optimal multipliers. Including more "noise" in the performance measures implies less precise inferences about the agent's action and, thus, a higher risk premium to induce a particular action.

Third, we examine settings in which only the variance of the weighted likelihood ratios matters (their mean is always zero), either because of particular agent preferences (square-root utility) or because the distribution of the likelihood ratios is fully characterized by their variance-covariance matrix (normal distributions). When comparing two performance evaluation systems, the variance of the weighted likelihood ratios is larger with one system than the other *independently of the multiplier* if, and only if, the corresponding variance-covariance matrix of the likelihood ratios is larger (i.e., their difference is positive definite). This is our VCM condition.

When the performance measures are both stochastically and technologically independent, i.e., each performance measure is affected by only one task, the variance-covariance matrix of the likelihood ratios is a diagonal matrix: the VCM condition requires that the variances of the likelihood ratios are larger with one information system than the other system for all tasks. When the performance measures are either technologically and/or stochastically dependent, the covariances of the likelihood ratios across the different tasks also enter the VCM condition; in these settings, the notion of more variability in the likelihood ratio and, thus, more precise inferences about the agent's multiple tasks, is a statement about the size of the full variance-covariance matrix of the likelihood ratios and not only about their variances.

We use the VCM dominance condition to derive results that hold for both optimal and LEN contracts. While this does not imply that LEN results have general validity, those that hinge on the VCM condition do. For example, our sufficient statistic result on the value of an additional performance measure generalizes Holmström's (1979) informativeness result to a multi-task setting with normally distributed performance measures and Feltham and Xie's (1994) LEN informativeness result to optimal contracts. The VCM condition allows a generalization of the conditional controllability principle to the multi-task setting and explains why the LEN results of Feltham and Xie (1994) point in the right direction. Other examples include the value of more precise information and linear aggregation of performance measures.

We find the concept of congruity too limited for ranking information systems over the set of implementable actions. Congruity can be defined only in LEN models on the equilibrium path in terms of the relative position of the first-best and second-best action choice vectors. The VCM condition ranks information systems for all implementable actions, not only on the equilibrium path. For both the LEN model and square-root utility with optimal contracts, the VCM condition also determines the set of implementable actions. By not restricting ourselves to the equilibrium

[2] As shown in the proof of Proposition 1, the optimal multiplier on the incentive compatibility constraints is the optimal multiplier that implements a given action when the principal randomizes over the two information systems. In the single-task setting, stochastic dominance of the distributions of the likelihood ratios alone is sufficient because any (scalar) multiplier "cancels out" (see Lemma A.1 in Appendix A).

path, our analysis of the value of information in a multi-task setting differs significantly from the accounting literature that followed Feltham and Xie (1994), such as Feltham and Wu (2000), Datar et al. (2001), or Demski et al. (2007).

Our analysis demonstrates that some LEN results, including many of the key results in Feltham and Xie (1994), are robust to removing the linear contract restriction and shows that the key assumption driving these results (via the VCM condition) is the linear structure of the performance measures with normally distributed additive noise independent of action choices.[3] We do not claim that results from LEN models are likely to be valid in most performance evaluation settings of interest.[4] Rather, we provide a new formulation of the LEN framework in terms of likelihood ratios as for optimal contracts, which can provide the basis for a deeper understanding of when and why LEN results may have more general validity.

II. THE GENERAL MULTI-TASK AGENCY MODEL

The model in this section is based on the single-period multi-task model of Holmström and Milgrom (1991) and Feltham and Xie (1994) (without the LEN restrictions we introduce in the next section). A risk-neutral principal owns a technology that requires productive effort on m tasks $a = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$ from a risk- and effort-averse agent.[5] The agent's preferences over compensation and effort are given by:

$$u(c,a) = u(c)k(a) - v(a), \tag{1}$$

where the utility of compensation, $u(c)$, is increasing and concave, and the effort cost is either multiplicatively separable ($v(a) = 0$) or additively separable ($k(a) = 1$). The agent's effort cost, $v(a)$ or $k(a)$, is increasing in all directions and convex.

The principal minimizes the expected cost of compensation $c(y)$ that induces the agent to choose a given action a. The contract $c(y)$ the principal offers the agent is a function of n performance measures $y \in \mathbb{R}^n$ whose distribution has joint density $\phi(y|a,\eta)$, where η is the information system that produces the contractible performance measures y. The contracts are bounded from below, $c(y) \geq \underline{c}$, to avoid the Mirrlees (1999) problem of non-existence of an optimal solution to the principal's problem. The agent accepts the contract if its expected utility is greater than or equal to his/her reservation utility (the agent's participation, or individual rationality constraint)

$$U^a(c,a|\eta) = \int u(c(y))k(a)\phi(y|a,\eta)dy - v(a) \geq \bar{U}. \tag{2}$$

The agent chooses action a to maximize his/her expected utility, which we assume is equivalent to a satisfying the first-order condition (the agent's incentive compatibility constraint):

---

[3] Holmström and Milgrom (1987) introduced the LEN framework in the agency literature as the limiting case of an optimal contract setting and the study has been much cited in support of using a LEN model in the early years after its publication. However, their approach has been shown to be fragile to changing any of the assumptions, see Lambert (2001), Christensen and Feltham (2005), or Hellwig and Schmidt (2002).

[4] If the agent's actions affect the level of risk in the performance measures, then optimal contracts shield the agent from low outcomes, which is not possible with linear contracts (Christensen and Feltham 2005, Ch. 21). In multiperiod settings, optimal full-commitment contracts exploit intertemporal correlation to a greater extent than can be achieved with linear contracts, while optimal renegotiation-proof contracts yield similar results to those obtained with renegotiation-proof linear contracts (Christensen and Feltham 2005, Chs. 27 and 28).

[5] In what follows, we use the following vector and matrix notation. Vectors are thought of as column vectors in all cases, and the scalar product of two vectors a and b in $\mathbb{R}^m$ is denoted by $a \cdot b$. For a matrix $M = [m_{ij}]_{1 \leq i \leq n, 1 \leq j \leq m}$ with n rows and m columns, we do not distinguish between the matrix and the associated linear operator $M: \mathbb{R}^m \to \mathbb{R}^n$ defined by $Mb = (\Sigma_{1 \leq j \leq m} m_{ij} b_j)_{1 \leq i \leq n}$. We denote the transpose matrix as M^T. Throughout, we use $A^T a \cdot b = a \cdot Ab$ and $(AB)^T = B^T A^T$.

$$\nabla_a U^a(c,a|\eta) = \int u(c(y))[\nabla_a k(a)\phi(y|a,\eta) + k(a)\nabla_a \phi(y|a,\eta)]dy - \nabla_a v(a) = 0, \quad (3)$$

where ∇_a denotes the gradient vector of partial derivatives $(\partial / \partial a_1, \partial / \partial a_2, \ldots, \partial / \partial a_m)$.

The optimal contract that implements a minimizes the agent's expected compensation subject to the participation and incentive compatibility constraints (2) and (3). The first-order condition with respect to $c(y)$ (when $c(y) > \underline{c}$) characterizes the contract as follows:

$$\frac{1}{u'(c(y))} = k(a)\lambda(a) + \mu(a) \cdot \nabla_a k(a) + k(a)\mu(a) \cdot L(\phi(y), \eta|a), \quad (4)$$

where $\lambda(a)$ and $\mu(a)$ are the multipliers for the participation and incentive compatibility constraints, respectively, and $L(\phi(y),\eta|a)$ is the likelihood ratio:

$$L(\phi(y), \eta|a) = \frac{\nabla_a \phi(y|a,\eta)}{\phi(y|a,\eta)}. \quad (5)$$

Kim (1995) shows that the efficiency of information systems for implementing a single-task action a can be compared using the second-order stochastic dominance (SSD) criterion for the distributions of the likelihood ratios (or, equivalently, by mean-preserving spreads using the fact that likelihood ratios have zero mean). In a single-task setting, the likelihood ratio is one-dimensional (even with multiple performance measures), while in a multi-task setting it is multi-dimensional and cannot be compared across information systems using the SSD criterion (which is a single-dimensional concept). However, the likelihood ratio enters the characterization of optimal contracts, see (4), as the single-dimensional weighted likelihood ratio $\mu(a)\cdot L(\phi,\eta|a)$. Our main result shows that the SSD condition on the weighted likelihood ratio is sufficient to rank information systems and, thus, generalizes Kim (1995, Proposition 1) to a multi-task setting.

Proposition 1: Assume the agent's utility is either additively or multiplicatively separable in compensation and effort, the agent's incentive compatibility constraints can be represented by their first-order conditions, and that a is an implementable action under two information systems η^1 and η^2. There exists a vector of multipliers $\mu(a)$ such that, if the distribution function of $\mu(a) \cdot L(\phi, \eta^1 \mid a)$ second-order stochastically dominates the distribution of $\mu(a) \cdot L(\phi, \eta^2 \mid a)$, then the principal prefers η^2 to η^1 for implementing action a.

All proofs are in Appendix A.

As shown in the proof, $\mu(a)$ is the optimal multiplier that implements a when the principal randomizes over the two information systems. In the single-task setting, stochastic dominance of the distributions of the likelihood ratios alone is sufficient because any (scalar) multiplier $\mu(a)$ "cancels out" (see Lemma A.1 in Appendix A).

We do not restrict the type of utility functions (other than effort cost separability in (1)) or the distributions of the performance measures.[6] Relying on the first-order approach restricts the settings for which Proposition 1 is applicable.[7] We show in Appendix B that the same SSD condition on the weighted likelihood ratios is sufficient to rank information systems in settings with finite

[6] Kim (1995) and Kim and Suh (1991) assume additively separable effort cost; we also allow multiplicatively separable effort cost because this is standard in the LEN models we examine later. We use integral notation throughout for simplicity; the proofs also hold for discrete distributions of the performance measures.

[7] There are no sufficient conditions reported in the literature for the applicability of the first-order approach in multi-task settings. Mirrlees (1999), Rogerson (1985), Jewitt (1988), Sinclair-Desgagné (1994, 2009), and Conlon (2009) derive sufficient conditions for the applicability of the first-order approach in single-task agencies.

action sets: Proposition B.1 is a discrete action version of Proposition 1 that does not rely on the first-order approach. With a finite action set, the ranking involves likelihood ratios for all the binding incentive compatibility constraints, while the first-order approach is akin to assuming that only the "adjacent" incentive constraints are binding.

We assume that the action a is implementable with either information system. In general, we cannot characterize the implementable actions, although with either square-root utility or under the LEN assumptions the problem becomes tractable through the explicit calculation of the multipliers on the first-order incentive compatibility constraints—see Proposition 4 and Lemma 1 below.[8]

The SSD condition in Proposition 1 depends on the endogenously determined optimal multiplier on the incentive compatibility constraints and, thus, on the specific characteristics of the incentive problem, such as agent preferences. In the following subsections, we identify conditions for ranking information systems independently of these characteristics.

Technological and Stochastic Independence

We begin with a setting in which the performance measures are technologically and stochastically independent; that is, the performance measures can be grouped into a set of (possibly multi-dimensional) task-specific performance measures $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$, such that:

$$\phi(y|a,\eta) = \phi(y_1|a,\eta) \ldots \phi(y_m|a,\eta) = \phi(y_1|a_1,\eta) \ldots \phi(y_m|a_m,\eta), \tag{6}$$

where the first factorization is stochastic independence and the second factorization is technological independence. Then, the likelihood ratios, $L_i(y_i \mid a_i, \eta) = \phi_{a_i}(y_i \mid a_i, \eta) / \phi(y_i \mid a_i, \eta)$, are stochastically independent, and the SSD condition in Proposition 1 is implied by SSD conditions for each task separately, independently of the multipliers; the strong separability of performance measures across tasks allows "cancellation" of the multipliers as in the single-task case.

Proposition 2: Assume the performance measures are technologically and stochastically independent and that a is an implementable action under two information systems η^1 and η^2. If the distribution function of $L_i(y_i \mid a_i, \eta^1)$ second-order stochastically dominates the distribution of $L_i(y_i \mid a_i, \eta^2)$ for all $i = 1, \ldots, m$, then the principal prefers η^2 to η^1 for implementing action a.

This result is parallel to the single-task result in Kim (1995, Proposition 1): if the performance measures are technologically and stochastically independent, then more variation in the likelihood ratio *for each task* implies that the risk premium required for inducing a particular multi-task action can be reduced.[9]

Blackwell Garbling

We next show that Blackwell garbling implies the SSD condition in Proposition 1 for any multiplier. For performance measures y_1 and y_2 associated with performance evaluation systems η^1 and η^2, we say that y_1 is a Blackwell garbling of y_2 if there exists a function $h(y_1, y_2) \geq 0$ such that the following three conditions hold (Kim 1995; Marschak and Miyasawa 1968, Condition B):

$$\phi(y_1|a) = \int h(y_1,y_2)\phi(y_2|a)dy_2 \tag{7a}$$

[8] Dewatripont et al. (1999) discuss the action implemented in equilibrium in a context with multiple tasks and implicit contracts based on career concerns. Note, however, the differences: we are interested in endogenous optimal explicit contracts, while they work with an implicit contract that is the expectation of an unobservable parameter conditional on observing performance measures.

[9] Proposition 2 also holds in the discrete-action setting of Appendix B when only the "adjacent" incentive constraints are binding.

$$\int h(y_1,y_2)dy_1 = 1 \quad \text{for all } y_2 \tag{7b}$$

$$0 < \int h(y_1,y_2)dy_2 < \infty \quad \text{for all } y_1. \tag{7c}$$

Proposition 3: If y_1 is a Blackwell garbling of y_2, then $\mu \cdot L(\phi, \eta^1 \mid a)$ second-order stochastically dominates $\mu \cdot L(\phi, \eta^2 \mid a)$ for all μ and all a. Thus, for any action a implementable under η^1 and η^2, the principal prefers η^2 to η^1 for implementing a.

Proposition 3 generalizes Kim (1995, Proposition 4) to our multi-task setting: including more "noise" in the performance measures implies less precise inferences about the agent's action and, thus, a higher risk premium to induce a particular action. In addition, the proof of Proposition 3 carries through without modifications to the setting with a finite action set in Proposition B.1, Appendix B.[10]

VCM Dominance

As demonstrated by Kim and Suh (1991) for the single-task setting, if either the agent's utility function is restricted or the performance measures have particular distributions, the ranking of information systems reduces to ranking the variances of the respective likelihood ratios. In a multi-task setting, under similar assumptions, the ranking is based on the variances of the weighted likelihood ratios, i.e., $\Sigma L(\phi,\eta|a)\mu\cdot\mu$, where $\Sigma L(\phi,\eta|a)$ is the variance-covariance matrix of the likelihood ratios. To obtain a ranking independent of the optimal multipliers on the incentive compatibility constraints, we rank the variance-covariance matrices of the likelihood ratios and define the VCM dominance condition as follows.

Definition 1 (VCM Dominance): Information system η^2 strictly dominates information system η^1 at a in the VCM (Variance-Covariance Matrix) sense if:

$$\Sigma L(\phi,\eta^2|a) > \Sigma L(\phi,\eta^1|a). \tag{8}$$

The strict inequality in the VCM condition (8) means that the difference between the two matrices is positive definite. A weak dominance relation is defined by a weak inequality in (8), i.e., the difference between the matrices is semi-positive definite.[11] The VCM condition gives only a partial ordering in the multi-task setting, because we are ranking multi-dimensional matrices. When the performance measures are technologically and stochastically independent, their variance-covariance matrix is diagonal and the VCM condition reduces to comparing the variances of the likelihood ratios for each task separately; otherwise, the VCM dominance condition must also include the covariances among the likelihood ratios.

---

[10] Kim (1995) points out that the Blackwell garbling condition is a strictly stronger condition than both ranking two information systems for all agency contexts (his Proposition 3) and for the SSD condition (his Proposition 5). His proofs rely on discrete action (and discrete performance measure) counter-examples; because our Proposition 3 also holds in these settings, Kim's (1995) analysis carries through in our multi-task setting.

[11] A square matrix A is positive (semi-positive) definite if $Ax\cdot x > 0$ ($Ax\cdot x \geq 0$), for all nonzero vectors x. In a sense, the VCM condition is too strong, since it ensures that $\Sigma L(\phi, \eta^2 \mid a)\mu \cdot \mu > \Sigma L(\phi, \eta^1 \mid a)\mu \cdot \mu$, for all $\mu \in \mathbb{R}^m$. For example, it is easy to show (using a similar argument as in Jewitt [1988, Lemma 1]) that if the agent has additively separable utility, then $\mu \cdot \nabla_a v(a) \geq 0$ (contrary to the single-task setting, this does not imply that all the multipliers are positive). Thus, the difference between the variance matrices needs to be positive-definite only on this half-space.

The next two results demonstrate the sufficiency of the VCM condition in ranking information systems if either (1) the agent's utility is additively separable with square-root utility of compensation or (2) the performance measures are normally distributed with a mean linear in agent effort and a variance independent of agent effort.

Proposition 4: Assume that the agent's utility is additively separable with square-root utility of compensation, i.e., $u(c, a) = 2\sqrt{c} - v(a)$, and that the optimal compensation is everywhere interior. For any implementable action a, if the variance-covariance matrices of the likelihood ratios satisfy $\Sigma L(\phi, \eta^2 \mid a) > \Sigma L(\phi, \eta^1 \mid a)$, then the principal strictly prefers information system η^2 to η^1 for implementing a. The set of implementable actions is $A(\eta) = \{a \in \mathbb{R}^m \mid \nabla_a v(a) \in \text{Im}(\Sigma L(\phi, \eta \mid a))\}$.

Proposition 4 generalizes Kim and Suh (1991, Proposition 1) to the multi-task setting.[12] The implementable actions are characterized by the row space of the variance-covariance matrix of the likelihood ratios, while in the single-task setting any action can be implemented as long as a nontrivial effort level can be induced.[13]

The following example illustrates how the implementable actions are constrained by the performance information available. Suppose the agent performs two tasks $a = (a_1, a_2)$ and there is only one performance measure available for contracting, $y = m_1 a_1 + m_2 a_2 + \varepsilon$. If $\psi(\varepsilon)$ is the probability density function of ε, then the probability density function of y is $\phi(y \mid a) = \psi(y - m_1 a_1 - m_2 a_2)$ and the likelihood ratio vector is:

$$L(\phi, y|a) = -(m_1, m_2)(\psi'/\psi)(y - m_1 a_1 - m_2 a_2) = -(m_1, m_2)L(\psi, \varepsilon).$$

The variance-covariance matrix of the likelihood ratios is singular because:

$$\Sigma L(\phi|a) = \begin{bmatrix} m_1^2 & m_1 m_2 \\ m_1 m_2 & m_2^2 \end{bmatrix} \text{var}(L(\psi, \varepsilon)) \tag{9}$$

and both rows are collinear with the vector (m_1, m_2); the row space of the variance-covariance matrix (9) is the single-dimensional space (line) determined by the vector (m_1, m_2). In other words, the only implementable actions are those for which the marginal effort cost $\nabla_a v(a)$ is collinear with (m_1, m_2). In particular, for a simple quadratic effort cost of the form $v(a) = \frac{1}{2} a \cdot a$, the implementable actions must satisfy $a_1 / a_2 = m_1 / m_2$.[14]

Next, assume the performance measures are joint normally distributed, with $y = Ma + \varepsilon$, where M is an $n \times m$ matrix of sensitivities of the n performance measures to the m tasks, $\varepsilon \sim N(0, \Sigma)$, and Σ is the positive definite $n \times n$ variance-covariance matrix of the n-dimensional performance measure noise vector ε. The m-dimensional likelihood ratio is then $L(\phi(y), \eta \mid a) = M^T \Sigma^{-1}(y - Ma)$, and its $m \times m$ variance-covariance matrix is $\Sigma L(\phi, \eta) = M^T \Sigma^{-1} M$.[15] Thus,

---

[12] The ranking of information systems also holds in the discrete action setting of Appendix B, if we replace the VCM condition with ranking the variances of the likelihood ratios in (B7). Proposition 4 is derived explicitly from the first-order condition for optimal contracts (4) with square-root utility, and not from the general SSD condition in Proposition 1. Even if $\Sigma L(\phi, \eta^2 \mid a) > \Sigma L(\phi, \eta^1 \mid a)$, it is not possible to establish the SSD condition in Proposition 1 for general distributions.

[13] The row (column) space of a matrix is the linear subspace of vectors that can be written as a linear combination of the row (column) vectors of the matrix. Similarly, the "span" of a set of vectors is the linear subspace of vectors that can be written as a linear combination of these vectors.

[14] Because the implementable actions have the same characterization, this example also works for LEN models, see Lemma 1 and Lemma 2 below, and the discussion following. For LEN models, this example is familiar from the discussion of noncongruity in Feltham and Xie (1994).

[15] The likelihood ratio depends on the action a, but is a linear function of the signal y and, thus, its variance-covariance

$\mu \cdot L(\phi, \eta|a)$ are scalar zero-mean normally distributed random variables and the SSD condition is equivalent to comparing their variances, $\Sigma L(\phi,\eta)\mu \cdot \mu$. If the VCM condition $\Sigma L(\phi, \eta^2) > \Sigma L(\phi, \eta^1)$ is satisfied, then $[\Sigma L(\phi, \eta^2) - \Sigma L(\phi, \eta^1)]\mu \cdot \mu > 0$ for all μ and the variance is higher for system η^2 than for system η^1. Hence, in this setting, the VCM condition implies the general SSD condition in Proposition 1 and is sufficient for ranking the two information systems.

Proposition 5: If the performance measures are normally distributed with $y \sim N(Ma,\Sigma)$ and if $\Sigma L(\phi, \eta^2) > \Sigma L(\phi, \eta^1)$, then the principal strictly prefers information system η^2 to information system η^1 for implementing a given action a. For a weak preference relation, one replaces the strict inequality by a weak inequality. In addition, if y_1 is a Blackwell garbling of y_2, then η^2 dominates η^1 in the VCM sense.

Proposition 5 generalizes Kim and Suh (1991, Proposition 2) to the multi-task setting. We cannot prove or disprove that Blackwell garbling is strictly stronger than the VCM condition. Blackwell garbling requires a kernel $h(y_1, y_2)$ that is independent of actions. Kim's (1995) examples, showing that neither the ranking of information systems in an agency context nor the SSD condition imply Blackwell garbling, are based on the local nature of the likelihood ratio. However, under the assumptions in Proposition 5, the VCM condition is independent of the action induced. This is because the agent's action does not affect the variance of the performance measures, the variance-covariance matrix of the likelihood ratio does not depend on a and the VCM condition is "global" with respect to a. As a result, Kim's (1995) counter-example (see his Proposition 3 and Proposition 5) that the SSD condition does not imply Blackwell garbling cannot be extended in our context to the VCM condition.

We also cannot prove or disprove that the variance of likelihood ratio condition is necessary for ranking two information systems in the agency model.[16] We conjecture that the ranking of information systems in all agencies does not imply the VCM condition. The VCM condition at a requires that for all $\mu \in \mathbb{R}^m$:

$$\Sigma L(\phi, \eta^2|a)\mu \cdot \mu \geq \Sigma L(\phi, \eta^1|a)\mu \cdot \mu. \tag{10}$$

By Proposition 1, when the performance measures are normally distributed, (10) is sufficient to rank the information systems for all implementable actions a if it satisfied for all optimal multipliers $\mu(a)$. However, if the optimal multipliers $\mu(a)$ for all implementable actions a do not span $\mathbb{R}^m$, even if the principal prefers η^2 over η^1 for all implementable a, we cannot expect to prove that (10) holds for all $\mu \in \mathbb{R}^m$. This is not an issue in the single-task case because $\Sigma L(\phi, \eta^2)\mu \cdot \mu \geq \Sigma L(\phi, \eta^1)\mu \cdot \mu$ for one nonzero multiplier μ is sufficient to prove that the scalar variances satisfy $\Sigma L(\phi, \eta^2) \geq \Sigma L(\phi, \eta^1)$—see Kim and Suh (1991, Proposition 2).

III. THE MULTI-TASK LEN MODEL

In this section, we introduce the LEN assumptions in the multi-task agency model. The agent has exponential utility of consumption $u(c) = -\exp(-r[c - \kappa(a)])$ with multiplicatively separable effort cost $\kappa(a)$, which is a convex function of the agent's action and is increasing in each task separately. The principal offers the agent a linear compensation contract $c(y) = f + v \cdot y$, where y is a vector of n performance measures reported by information system η. As in the setting of Proposition 5, the performance measures are joint normally distributed with $y \sim N(Ma,\Sigma)$.

---

matrix does not depend on the agent's action. This is reflected in the notation $L(\phi,\eta|a)$ for the likelihood ratio and $\Sigma L(\phi,\eta)$ for its variance-covariance matrix.

[16] We have this result for the "standard" LEN model: in Proposition 7 below, the ranking of information systems in all agencies on the equilibrium path implies the VCM condition.

The LEN assumptions imply the agent's compensation is normally distributed and the agent's preferences have a mean-variance representation given by the certainty equivalent:

$$\mathrm{CE}(c(y)|a,\eta) = \mathrm{E}[c(y)|a,\eta] - \frac{1}{2}r\mathrm{var}(c(y)|\eta) - \kappa(a). \tag{11}$$

The principal seeks to induce an action a with a linear contract $c(y)$ at minimum expected cost. The agent accepts the contract if its certainty equivalent is greater than or equal to his/her reservation certainty equivalent CE_0 (the agent's participation, or individual rationality constraint):

$$\mathrm{CE}(c(y)|a,\eta) \geq \mathrm{CE}_0. \tag{12}$$

The agent's action choice maximizes his/her certainty equivalent of compensation net of effort costs, which in this case is equivalent to a satisfying the first-order (incentive compatibility) constraint:

$$\nabla_a \mathrm{CE}(c(y)|a,\eta) = \nabla_a \mathrm{E}[c(y)|a,\eta] - \nabla_a \kappa(a) = 0, \tag{13}$$

where the agent's action does not impact the variance of compensation because it does not impact the variance of y and the contracts are linear.[17] The principal's problem is to minimize expected compensation cost over the vector space of linear contracts C, $\min_{c(y)\in C}\mathrm{E}[c(y) \mid a, \eta]$, subject to the individual rationality constraint (12) and the incentive compatibility constraint (13) for a given action a.

Lemma 1: Under the assumptions of the LEN model:

(i) For an implementable action a, an optimal multiplier on the incentive compatibility constraint (13) is denoted by $\mu = r\Sigma L(\phi, \eta)^{-1}\nabla_a \kappa(a)$ and is characterized by:

$$\Sigma L(\phi,\eta)\frac{\mu}{r} = \nabla_a \kappa(a). \tag{14}$$

(ii) The set of implementable actions $A(\eta)$ is characterized by the row space of M and the cost of effort $\kappa(a)$, $A(\eta) = \{a \in \mathbb{R}^m \mid \nabla_a \kappa(a) \in \mathrm{Im}(M^T)\}$.

(iii) For any $a \in A(\eta)$, the optimal linear contract that implements a is characterized by:

$$c(y) = \mathrm{CE}_0 + k(a) + \frac{1}{2}r\Sigma L(\phi,\eta)^{-1}\nabla_a \kappa(a)\cdot\nabla_a \kappa(a) + \Sigma L(\phi,\eta)^{-1}\nabla_a \kappa(a)\cdot L(\phi,\eta|a). \tag{15}$$

The contract is uniquely determined, even when the multiplier μ that satisfies (14) is not.

Lemma 1 recasts the LEN model in a new form. Using the certainty equivalent representation of the agent's utility, we rewrite the principal's problem to include the mean and variance of the agent's compensation; this leads to optimal linear contracts that depend on the likelihood ratio and on multipliers that can be explicitly characterized.

Action a is implementable if, and only if, the marginal effort cost vector $\nabla_a \kappa(a)$ is in the image of $\Sigma L(\phi,\eta)$, which is equivalent to $\nabla_a \kappa(a)$ being in the row space of the sensitivity matrix M. For example, if the agent's effort cost is of the simple quadratic form $Ka\cdot a$, where K is a positive definite matrix, and $\Sigma L(\phi,\eta)$ is non-singular, or equivalently M is of rank m, then any

[17] We allow for the possibility that the first-order condition (13) does not uniquely determine the agent's action choice. For example, if $\kappa(a) = \kappa(p\cdot a)$, then $\nabla_a \kappa(p\cdot a) = p\kappa'(p\cdot a)$, and the agent is indifferent between vectors of tasks for which $p\cdot a$ is the same. In this case, we assume that the principal can specify the agent's action.

action $a \in \mathbb{R}^m$ is implementable. For an effort cost function of the form $\kappa(p \cdot a)$, all actions are implementable if p is in the row space of M, and the principal specifies the agent's action among those over which the agent is indifferent. If p is not in the row space of M, then there is no interior solution for the agent's action choice.[18]

An optimal LEN contract pays the agent a fixed wage consisting of his/her reservation certainty equivalent, the cost of effort, the risk premium, and a linear mean zero variable wage. Both the risk premium and the variable wage depend on the multiplier μ on the agent's incentive compatibility constraint, which in turn depends on the action a to be implemented. The LEN model allows an explicit characterization of this multiplier, although it may not be uniquely determined if $\Sigma L(\phi, \eta)$ is singular. However, as we show in the proof of Lemma 1, both $\mu \cdot \nabla_a \kappa(a)$ and $\mu \cdot L(\phi, \eta \mid a)$ are uniquely determined; this in turn allows an explicit calculation of the expected cost of implementing action a with an optimal linear contract:

$$\mathrm{E}[c(y)|a, \eta] = \mathrm{CE}_0 + \kappa(a) + \frac{1}{2} r \Sigma L(\phi, \eta)^{-1} \nabla_a \kappa(a) \cdot \nabla_a \kappa(a). \tag{16}$$

Because the principal prefers the information system that allows implementing a with a lower risk premium, and this is determined by the variance-covariance matrix of the likelihood ratio, we are led again to the VCM condition.

To compare the efficiency of two information systems in implementing a given action, we need to consider actions that are implementable under both. The VCM condition ranks the sets of implementable actions.

Lemma 2: If $\Sigma L(\phi, \eta^2) \geq \Sigma L(\phi, \eta^1)$, then $A(\eta^2) \supseteq A(\eta^1)$.

In other words, if η^2 dominates η^1 in the VCM sense, the principal can implement at least as many actions with η^2 as with η^1; in particular, the actions implementable under both η^1 and η^2 are the same as the actions implementable under η^1.

The next result demonstrates the sufficiency of the VCM condition for ranking information systems in LEN models when the same actions can be implemented with both.

Proposition 6: Under the LEN assumptions for two information systems η^1 and η^2, if the row spaces of M_1 and M_2 are equal, then the strict VCM condition $\Sigma L(\phi, \eta^2) > \Sigma L(\phi, \eta^1)$ is sufficient for the principal to strictly prefer η^2 to η^1 for implementing any action $a \in A(\eta^1) = A(\eta^2)$. The weak VCM condition is sufficient for a weak preference of η^2 over η^1.

The VCM condition is sufficient for ranking information systems in the LEN case because the optimal linear contracts depend on the same linear function of the likelihood ratio as the general optimal contracts. The VCM condition in Proposition 6 does not depend on particular functional forms for the principal's expected benefit and for the agent's action cost $\kappa(a)$. Thus, we rank performance measures in any agency for all implementable actions and not only for the optimal action in a particular parametrization of the model. However, the set of implementable actions $A(\eta^1) \subseteq A(\eta^2)$ may depend on the agent's effort cost function $\kappa(a)$: if either M_1 or M_2 are not of full rank m, the set of implementable actions is limited and $\nabla_a \kappa(a)$ enters its characterization (see also (A28) in Appendix A and the subsequent discussion).

[18] As a simple example, consider the case of a single performance measure and two tasks specified as $y = m_1 a_1 + m_2 a_2 + \varepsilon$, $\kappa(a_1, a_2) = \kappa(a_1 + a_2)$, and $m_1 > m_2$. For any contract $c = f + vy$, if $v > 0$, and a_1, a_2 are not exogenously bounded, the agent can choose a_1 arbitrarily large, $a_2 = -a_1$, and derive an unboundedly large benefit $v(m_1 - m_2)a_1$ while keeping effort costs constant at $\kappa(0)$. This leads to a contract without incentives, $v = 0$ as in Holmström and Milgrom (1991), or to boundary solutions determined by constraints on a_1, a_2.

We have assumed that the same actions are implementable under both information systems. When $\Sigma L(\phi, \eta^1)$ is non-singular and, thus, has full rank m, all actions are implementable under both information systems because both variance-covariance matrices have full rank. Otherwise, when $\Sigma L(\phi, \eta^1)$ is singular and not all actions are implementable, if the row spaces of M_1 and M_2 are equal, then the same actions are implementable with both systems. If η^2 allows implementation of a larger set of actions than η^1, then Proposition 6 is still applicable with some minor modifications. Specifically, if the VCM condition is satisfied on a linear subspace S in the intersection of the row spaces of M_1 and M_2, and if S is invariant to both variance-covariance matrices $\Sigma L(\phi, \eta^i)$, then the principal prefers η^2 to η^1 for implementing any action, such that $\nabla_a \kappa(a) \in S$.

The "Standard" Multi-Task LEN Model

Feltham and Xie (1994) use a particular parametrization of the principal's benefit function and the agent's effort cost that allows an explicit characterization of the optimal action to be induced in equilibrium. We re-examine their "standard" formulation of the LEN model and show that the VCM condition is both necessary and sufficient to rank information systems on the equilibrium path. As in Feltham and Xie (1994), we assume the principal's expected benefit is $b(a) = b \cdot a$, where $b = (b_1, \ldots, b_m)$ are the marginal benefits from each of the agent's tasks, and $\kappa(a) = \frac{1}{2} a \cdot a$. The principal maximizes his/her expected benefit net of compensation costs, $U^p(a, \eta) = b \cdot a - \mathrm{E}[c(y) \mid a, \eta]$, by choosing the optimal action to be induced in equilibrium with an optimal linear contract as determined in Lemma 1. The next result characterizes the equilibrium.

Lemma 3: In equilibrium, the optimal action induced is:

$$a = (\Sigma L(\phi, \eta) + rI)^{-1} \Sigma L(\phi, \eta) b,$$

and the optimal linear contract and the principal's optimal expected surplus are:

$$c(y) = \mathrm{CE}_0 + \frac{1}{2}(\Sigma L(\phi, \eta) + rI)^{-1} \Sigma L(\phi, \eta) b \cdot b + (\Sigma L(\phi, \eta) + rI)^{-1} b \cdot L(\phi, \eta | a) \tag{17}$$

$$U^p = \frac{1}{2}(I - r(\Sigma L(\phi, \eta) + rI)^{-1}) b \cdot b - \mathrm{CE}_0. \tag{18}$$

In Lemma 1 we showed that the optimal linear contract is linear in the likelihood ratio, and the LEN assumptions allow us to explicitly characterize the multipliers λ and μ. Lemma 3 determines all the equilibrium parameters—action, contract, and payoffs—in terms of the variance-covariance matrix of the likelihood ratio $\Sigma L(\phi, \eta)$; in the parametrization of the "standard" LEN model, the optimal multiplier μ and the optimal action in equilibrium depend on $\Sigma L(\phi, \eta)$.[19] Using the explicit expression (18) for the principal's surplus in equilibrium, we can rank information systems for inducing the optimal action for an arbitrary marginal benefit vector b.

Proposition 7: The principal strictly prefers η^2 to η^1 for implementing the optimal action for all benefit vectors b, if, and only if, η^2 strictly dominates η^1 in the VCM

[19] To reconcile the expressions for the principal's utility in (18) from Lemma 3, (A40) in Appendix A, and those in Feltham and Xie (1994), it is sufficient and straightforward to show that:

$$\Sigma L(\phi, \eta)(\Sigma L(\phi, \eta) + rI)^{-1} = M^T \Sigma^{-1} M (M^T \Sigma^{-1} M + rI)^{-1} = M^T (M M^T + r\Sigma)^{-1} M.$$

sense, $\Sigma L(\phi, \eta^2) > \Sigma L(\phi, \eta^1)$. A weak preference relation in all agencies is equivalent to the weak VCM condition.

The sufficiency of the VCM condition follows directly from the explicit expression of the principal's expected utility (18). To see that the VCM condition in Proposition 7 is also necessary if the ranking of information systems holds on the equilibrium path for all $b \in \mathbb{R}^m$, note that $(\Sigma L(\phi, \eta^2) - \Sigma L(\phi, \eta^1))b \cdot b \geq 0$ for all b implies $\Sigma L(\phi, \eta^2) - \Sigma L(\phi, \eta^1) \geq 0$. Two information systems are equivalent if they have the same variance-covariance matrix of the likelihood ratio.

For the standard LEN model with a quadratic effort cost, the set of implementable actions is the column space of the variance-covariance matrix of the likelihood ratio, which is the same as the row space of the sensitivity matrix M (this follows from (14) by using $\nabla_a \kappa(a) = a$); in our notation, $A(\eta) = \text{Im}(\Sigma L(\phi, \eta)) = \text{Im}(M^T)$. By Lemma 2, if the VCM dominance condition in Proposition 7 is satisfied, then the space of implementable actions under η^2 is at least as large as that under η^1. By Proposition 6, any action that is implementable under η^1 is implementable under η^2 at the same or at a lower cost. Thus, it is not surprising that the VCM condition is sufficient to rank information systems on the equilibrium path.

The ranking is valid, even if the variance-covariance matrices of the likelihood ratios are singular and the row spaces of M_1 and M_2 are different (i.e., the sets of implementable actions differ under the two information systems). As opposed to Proposition 6, where we rank the optimal linear contracts for the same actions that are implementable under both information systems, in Proposition 7, we rank the information systems on the equilibrium path for the endogenously determined optimal actions (which generally differ for different information systems).

Although the VCM condition is shown to be sufficient for ranking information systems in Proposition 6, it does not follow that this condition is necessary, even if the principal prefers one information system over the other for inducing any implementable action. For example, under the conditions for which we proved Proposition 6, if $\Sigma L(\phi, \eta)$ is non-singular, but $\kappa(a) = \kappa(p \cdot a)$, any action is implementable with multipliers that are all proportional to the single vector $\Sigma L(\phi, \eta)^{-1} p$ (also footnote 17). It follows that the principal prefers η^2 to η^1 for implementing any action $a \in \mathbb{R}^m$ if, and only if, $\Sigma L(\phi, \eta^1)^{-1} p \cdot p > \Sigma L(\phi, \eta^2)^{-1} p \cdot p$. Being satisfied for a single vector, this does not rank the variance-covariance matrices and, thus, does not imply the VCM condition.

There is no notion of "congruity" here. In Feltham and Xie (1994) and Datar et al. (2001), congruity is defined relative to $b \in \mathbb{R}^m$. As Demski et al. (2007) show, congruity can be interpreted in terms of the relative positions of the benefit vector b and the row space of M (which is the set of implementable actions). The VCM condition holds uniformly with respect to b and only determines how large is the set of implementable actions. Improvements in congruity, when the set of implementable actions does not change, are not reflected in the variance-covariance of the likelihood function, because such improvements are about the position of a given b relative to M.

IV. APPLICATIONS OF THE VCM CONDITION

In this section we present a series of results that rely on the sufficiency of the VCM condition for ranking information systems for the three previously considered particular cases with normally distributed performance measures: (1) optimal contracts that implement a given action, as in Proposition 5; (2) optimal LEN contracts that implement a given action, as in Proposition 6; and (3) optimal LEN contracts that implement the optimal action in equilibrium, as in Proposition 7.

Value of Additional Information

Consider two joint normally distributed performance measures y_1, y_2, where $y_i = M_i a + \varepsilon_i$, and $(\varepsilon_1, \varepsilon_2) \sim N(0, \Sigma)$ with joint variance-covariance partitioned as:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}. \tag{19}$$

We say that y_2 has no incremental value over y_1 if the principal is indifferent between the information systems $\{y_1\}$ and $\{y_1, y_2\}$, either with respect to implementing any action that is feasible under both information systems or for implementing the optimal actions. Our next result characterizes several equivalent conditions for y_2 to have no incremental value over y_1.

Proposition 8: If the performance measures y_1 and y_2 are joint normally distributed as above, then the following conditions are equivalent:

(i) y_2 has no incremental value over y_1 in implementing a given action with either optimal or LEN contracts, or in implementing the optimal action with LEN contracts;
(ii) there exists a matrix Q such that $M_2 = QM_1$ and $\Sigma_{21} = Q\Sigma_{11}$;
(iii) there exists a matrix Q such that $y_2 = Qy_1 + \varepsilon$, where $\varepsilon \sim N(0, \Sigma_\varepsilon)$, and y_1 and ε are independent;
(iv) y_1 is a sufficient statistic for (y_1, y_2) for all a: there exists a function $h(y_1, y_2)$ such that $\phi(y_1, y_2 \mid a) = h(y_1, y_2)\phi(y_1 \mid a)$, for all y_1, y_2 and all a;
(v) $\Sigma L(\phi, y_1, y_2) = \Sigma L(\phi, y_1)$;

In addition, any of the equivalent conditions (i)–(v) imply that y_2 is a Blackwell garbling of y_1.[20]

Conditions (ii)–(v) are purely statistical properties of the performance measures, one of them being that y_1 is a sufficient statistic for (y_1, y_2). The key link between condition (v) and the ranking of inclusive information systems in (i) is provided by the VCM condition that underlies our three ranking results for normally distributed performance measures: Proposition 5 on implementing a given action with optimal contracts; Proposition 6 on implementing a given action with LEN contracts; and Proposition 7 on implementing the optimal action with LEN contracts.

Condition (v) in Proposition 8 is the same as the weak mean-preserving spread (MPS) condition of Kim (1995) because we have normally distributed performance measures. In our case, the cumulative distribution functions of the likelihood ratios are mean-preserving spreads of each other, either for all a or for no a because the agent's effort does not affect the variance of the performance measures. As a result, the equivalence between condition (v) and the sufficient statistic condition (iv) is similar to Kim (1995, Proposition 2). Condition (ii) in Proposition 8 is the same as condition (13) in Feltham and Xie (1994) and is equivalent to condition (iii) that one performance measure is a noisy linear function of the other.

As in Holmström (1979) and Feltham and Xie (1994), we can use the sufficient statistic condition to determine whether an additional signal is incrementally valuable in a multi-task agency. Proposition 8 generalizes Holmström (1979, Proposition 3) to a multi-task agency with normally distributed performance measures whose mean is linearly affected by effort. That is, an additional performance measure y_2 is of no value if, and only if, y_1 is a sufficient statistic for the pair (y_1, y_2). Our result also incorporates Feltham and Xie (1994, Proposition 3) on the value of additional information for LEN models *on the equilibrium path*. Feltham and Xie's (1994) proof relies on the "standard" LEN formulation and direct calculation of payoffs. We rely instead on the VCM condition to rank information systems when the performance measures are normally distributed. Thus, we extend Feltham and Xie (1994, Proposition 3) to optimal linear contracts that implement any given action in a more general formulation of the LEN model.

[20] The Blackwell garbling condition does not imply any of the equivalent conditions (i)–(v). If y_1 is a Blackwell garbling of y_2, then, by Proposition 5, y_2 weakly dominates y_1 in the VCM sense. However, if y_2 weakly dominates y_1 in the VCM sense, it does not follow that y_1 has no incremental value over y_2.

The linear structure of the normally distributed performance measures not only guarantees the sufficiency of the VCM condition for ranking information, but also ensures the condition is independent of the agent's actions. This is important in that it avoids the problems associated with the "all a or no a condition." Amershi et al. (1990) show that the all a or no a condition must hold when the sufficient statistic condition—which is stated for all a—is necessary for additional information to have no value.[21] The all a or no a condition requires that the information aggregate $\mu(a)\cdot L(\phi,\eta|a)$ that is sufficient to implement a—and is generally local with respect to a—matches the global nature of the sufficient statistic condition with respect to a. The optimal contract that implements a can be written on a sufficient statistic with respect to all a if, and only if, the probability distributions of performance measures satisfy the all a or no a condition.

In general, even in a single-task agency, the necessity of the equivalent conditions (ii)–(v) for additional information to have no value does not hold. Which information system is preferred also depends on which action the principal wants to induce. For a multi-task agency, $\mu(a)$ is an m-dimensional vector and depends on the implemented actions (for example, in the LEN case, $\Sigma L(\phi, \eta)\mu(a) / r = \nabla_a \kappa(a)$, see (14)). Thus, even if the information content of the likelihood ratio is independent of a, the information content of $\mu(a)\cdot L(\phi,\eta|a)$ changes with a.[22] For this reason, the SSD condition on $\mu(a)\cdot L(\phi,\eta|\alpha)$ in Proposition 1 is unlikely to be necessary for ranking information systems in general. For the linear structure of normally distributed performance measures assumed here, the SSD condition can be replaced with the stronger VCM condition that applies for all μ and does not depend on a. The VCM condition has the global nature to evaluate the information content of performance measures and allows the equivalent conditions (ii)–(v) to be necessary for additional information to have no value.

Conditional Controllability

The controllability principle is fundamental in performance evaluation: a manager should only be held accountable for that which s/he can control. A key implication of Holmström's (1979) results on the value of an additional performance measure is that even though an agent cannot affect a specific performance measure, this measure may well enter the agent's optimal contract if it is correlated with uncontrollable events affecting other primary performance measures. This is Antle and Demski's (1988) *conditional controllability* principle: an additional performance measure is valuable if it is controllable by the agent conditional on other available performance information.

Consider two sets of performance measures $\eta^1 = \{y_1\}$ and $\eta^2 = \{y_1, y_2\}$ as in the preceding subsection. Define $\zeta = \{z\}$, where z is the "filtered" measure $z = y_2 - \mathrm{E}[y_2 \mid y_1, a] = y_2 - [M_2 a + \Sigma_{21}\Sigma_{11}^{-1}(y_1 - M_1 a)]$. The joint density of the two performance measures can be written as $\phi(y_1, z \mid a) = \phi(y_1 \mid a)\phi(z \mid a)$. The likelihood ratio for the joint distribution is:

$$L(\phi, y_1, z|a) = \frac{\nabla_a \phi(y_1, z|a)}{\phi(y_1, z|a)} = \frac{\nabla_a \phi(y_1|a)}{\phi(y_1|a)} + \frac{\nabla_a \phi(z|a)}{\phi(z|a)} = L(\phi, \eta^1|a) + L(\phi, \zeta|a). \tag{20}$$

The two likelihood ratios are stochastically independent because the performance measures y_1 and z are stochastically independent. It follows that:

---

[21] The all a or no a condition holds if the likelihood ratio measure $\mu(a)\cdot L(\phi,\eta|a)$ has the same information content for all $(a,\mu(a))$ pairs, where $\mu(a)$ is the optimal multiplier in the optimal contract that implements a. That is, for any action to be implemented, $\hat{a}$, and for all other actions a, there exists a function ψ_a such that $\psi_a(\mu(\hat{a}) \cdot L(\phi, \hat{a})) = \mu(a) \cdot L(\phi, \eta \mid a)$.

[22] The all a or no a condition is also the reason why the relative incentive rates in the optimal linear aggregate of Banker and Datar (1989) are independent of a in the single-task case, but vary with a in the multi-task case.

$$\Sigma L(\phi, \eta^1, \zeta) = \Sigma L(\phi, \eta^1) + \Sigma L(\phi, \zeta) \geq \Sigma L(\phi, \eta^1). \tag{21}$$

In other words, the filtered performance measures $\zeta = \{z\}$ are incrementally valuable if they are controllable. Note that (η^1, ζ) is equivalent to η^2. In addition, z is controllable ($\nabla_a \phi(z \mid a) / \phi(z \mid a) \neq 0$ on a set of realizations of z of positive measure), if, and only if, y_2 is conditionally controllable ($\nabla_a \phi(y_2 \mid y_1, a) / \phi(y_2 \mid y_1, a) \neq 0$ on a set of realizations of (y_1, y_2) of positive measure). Thus, the additional performance measures y_2 are valuable if they are conditionally controllable.

This, of course, is Antle and Demski's (1988) conditional controllability principle for the special case of normal distributions, and it can also be found in Feltham and Xie (1994, Proposition 4) for the special case of a multi-task LEN model. We generalize Antle and Demski (1988) to the multi-task case and Feltham and Xie (1994, Proposition 4) to the optimal contract case. The same proof works for both linear and optimal contracts because it relies on the VCM condition.

If y_2 is not controllable, then $y_2 = M_2 a + \varepsilon_2$ with $M_2 = 0$, y_2 may still be conditionally controllable if ε_2 is correlated with ε_1. The variance-covariance matrix of the likelihood ratio with η^2 is:

$$\Sigma L(\phi, \eta^2) = [M_1^T \quad 0] \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} M_1 \\ 0 \end{bmatrix}. \tag{22}$$

The column space of $\Sigma L(\phi, \eta^2)$ is the same as the column space of $\Sigma L(\phi, \eta^1)$ and equal to the row space of M_1. By Proposition 4 and Lemma 1, the sets of implementable actions with square-root utility and in the LEN model are the same for η^1 and η^2, $A(\eta^2) = A(\eta^1)$. Thus, the uncontrollable performances measures y_2 do not expand the set of implementable actions, but they can be used to reduce the risk premium to induce a particular action if $\Sigma_{21} \neq 0$.

Value of More Precise Information

Assume there are two information systems η^1 and η^2 with the same number of performance measures and the same sensitivity to effort M, where $y_i = Ma + \varepsilon_i$ and $\varepsilon \sim N(0, \Sigma_i)$, such that one is "noisier" than the other in the sense that $\Sigma(\eta^1) > \Sigma(\eta^2)$. The VCM condition holds because $\Sigma L(\phi, \eta^i) = M^T \Sigma_i^{-1} M$, $\Sigma_2^{-1} > \Sigma_1^{-1}$, and for any vector $\mu \in \mathbb{R}^m$ we have:

$$\begin{aligned} \Sigma L(\phi, \eta^2)\mu \cdot \mu &= M^T \Sigma_2^{-1} M\mu \cdot \mu = \Sigma_2^{-1} M\mu \cdot M\mu > \Sigma_1^{-1} M\mu \cdot M\mu = M^T \Sigma_1^{-1} M\mu \cdot \mu \\ &= \Sigma L(\phi, \eta^1)\mu \cdot \mu. \end{aligned} \tag{23}$$

In other words, the less noisy system, that is η^2, is preferred for inducing any action a.

Informativeness and Linear Aggregation in Single-Task LEN Models

We now show that the optimal linear aggregate of two signals maximizes the variance of the likelihood ratio, which is the informativeness measure in the single-task case. This links the linear aggregation results of Banker and Datar (1989) to the informativeness results of Kim and Suh (1991) and explains why we get the "correct" linear aggregation with LEN models.

A direct calculation shows that, in the single-task, single performance measure "standard" LEN model with $y = ma + \varepsilon$, $\varepsilon \sim N(0, \sigma^2)$, the equilibrium action, contract, and payoffs can all be expressed in terms of m^2 / σ^2. This suggests itself as an informativeness measure for the performance measure y. The likelihood ratio is $L(\phi(y), \eta \mid a) = m\sigma^{-1}(y - ma)$ and its variance is m^2 / σ^2, which shows that m^2 / σ^2 is the right informativeness measure.

Consider two performance measures $y_i = m_i a + \varepsilon_i$, $i = 1,2$, where $(\varepsilon_1, \varepsilon_2) \sim N(0, \Sigma)$, $\text{var}(\varepsilon_i) = \sigma_i^2$, and $\text{cov}(\varepsilon_1, \varepsilon_2) = \zeta\sigma_1\sigma_2$. Then, the likelihood ratio of (y_1, y_2) and its variance are:

$$L(y_1,y_2|a) = \frac{1}{1-\zeta^2}\left[\left(\frac{m_1}{\sigma_1^2} - \frac{\zeta m_2}{\sigma_1\sigma_2}\right)(y_1 - m_1 a) + \left(\frac{m_2}{\sigma_2^2} - \frac{\zeta m_1}{\sigma_2\sigma_1}\right)(y_2 - m_2 a)\right]$$

$$\Sigma L(y_1,y_2) = \frac{1}{1-\zeta^2}\left[\frac{m_1^2}{\sigma_1^2} - 2\zeta\frac{m_1}{\sigma_1}\frac{m_2}{\sigma_2} + \frac{m_2^2}{\sigma_2^2}\right]. \quad (24)$$

The likelihood ratio and its variance for any linear aggregate $y = \alpha y_1 + \beta y_2$ are:

$$L(y|a) = \frac{\alpha m_1 + \beta m_2}{(\alpha^2\sigma_1^2 + 2\zeta\alpha\beta\sigma_1\sigma_2 + \beta^2\sigma_2^2)^{1/2}}(y - (\alpha m_1 + \beta m_2)a)$$

$$\Sigma L(y) = \frac{(\alpha m_1 + \beta m_2)^2}{\alpha^2\sigma_1^2 + 2\zeta\alpha\beta\sigma_1\sigma_2 + \beta^2\sigma_2^2}. \quad (25)$$

From the above, it is straightforward to show that aggregation is generally not desirable and that optimizing the variance of the likelihood ratio of the aggregate system in (25) gives an optimal linear aggregate with the same variance of the likelihood function as in (24). The relative incentive weights in the optimal linear aggregate of the two signals are $\alpha^\dagger / \beta^\dagger = (m_1\sigma_2^2 - \zeta m_2\sigma_1\sigma_2) / (m_2\sigma_1^2 - \zeta m_1\sigma_1\sigma_2)$. Thus, there is no loss from aggregation if, and only if, it is the optimal linear aggregation characterized in Banker and Datar (1989).

The linear aggregation of Banker and Datar (1989) holds for the LEN model, because the contracts are linear in the likelihood ratios. While any linear contract by definition linearly aggregates the performance measures, an optimal LEN contract is linear in the likelihood ratio, and the optimal linear aggregate is determined by the likelihood ratios as with optimal contracts.

V. CONCLUSIONS

Comparing information systems in agencies is the basis for managerial performance evaluation purposes. It has been long recognized that managers attend to several tasks, but we know how to compare information systems only in single-task agencies, such as in Holmström (1979) and Kim (1995). Most of what we know about the value of information in a multi-task setting comes from the multi-task LEN models of Holmström and Milgrom (1991), Feltham and Xie (1994), and Datar et al. (2001). While LEN provides much tractability to the multi-task environment, questions remain as to the robustness of the results given the linear contract restriction.

We derive information ranking results for multi-task models with both optimal and LEN contracts. For optimal contracts, we generalize the second-order stochastic dominance (SSD) condition in Kim (1995) by showing that it is sufficient to hold for the likelihood ratio weighted by the optimal multiplier on the incentive compatibility constraints. In general, Blackwell garbling implies our version of the SSD condition. In particular cases (square-root agent utility or normally distributed performance measures), the SSD condition can be replaced by the simpler variance-covariance matrix of likelihood ratios (VCM) condition that requires the ranking of the variance-covariance matrices of the likelihood ratios. The partial ordering of information systems by the VCM condition is the multi-task generalization of the complete ordering by the variance of the likelihood ratio in Kim and Suh (1991). By recasting the LEN model using likelihood ratios, we find that the same VCM condition is sufficient to rank information systems with LEN models. We use the VCM condition, for normally distributed performance measures, to prove a series of results that hold both with and without the LEN assumptions: the sufficient statistic results of Holmström (1979) and Feltham and Xie (1994) and, thus, the conditional controllability principle; the value of more precise information; and the optimal linear aggregation of performance information.

While we use the VCM condition in many of our results, we do not claim it has general applicability in ranking performance measures. For example, in Feltham and Wu (2001), the agent's effort affects the mean and the variance of a normally distributed performance measure (the standard deviation is a linear function of the agent's effort). In this case, the likelihood ratio is a quadratic function of the performance measure and, thus, it has a noncentral Chi-square distribution and the VCM condition no longer applies (unless the agent has square-root utility).[23]

LEN models have been appropriately criticized on the grounds that suboptimal use of information casts doubts on the results: "The restriction to linear contracts has been a controversial one in the agency literature. In particular, agency theory's intellectual roots lie in information economics, where information systems are compared based on the *optimal* use of information generated by the system. Restricting the contract to be linear is a very significant philosophical departure" Lambert (2001, 29; emphasis in the original). What, then, is the use of LEN models?

Our derivation of the optimal LEN contracts in terms of likelihood ratios and the validity of the VCM condition for both optimal contract and LEN models contributes to our understanding of the limits of using LEN models. The same VCM condition ranks information systems in three particular cases: square-root utility, normally distributed performance measures, and LEN models. However, this does not imply that the ranking of information systems in all three cases is the same. In other words, preferences over performance measures derived from one model, say LEN, do not necessarily hold for other models and *vice versa*. Because any result on the value of performance measures relying on the VCM condition will hold in all three settings, we can shed some light on *when* conclusions derived from a LEN-based analysis have more general validity.

The main justification for the use of LEN models remains their tractability. Holmström and Milgrom (1991, 24) point out that the general agency models "have remained too intractable to effectively address broader organizational issues such as asset ownership, job design, and allocation of authority," but at the same time "because the model is so tractable, we shall not avoid the temptation to stretch its use somewhat further in this article," Holmström and Milgrom (1991, 31). Tractability and the associated explicit calculations, if pushed too far, can obscure the information economics of the problem. Our derivation of the VCM condition for LEN models avoids some of the temptations of tractability: we found the "right" result by deliberately staying away from the usual calculations along the equilibrium path.

Hemmer (2004, 150) goes further in criticizing LEN models: "[t]he problem is that it is impossible to know whether the results provided by the LEN framework are actually consistent with the analysis of optimal contracts *unless* one does the latter part as well. But since the LEN approach often is used where the analysis of optimal contracts seems too complicated, such verification is rarely, if ever, provided. Moreover, if the general analysis could actually be done, there would be no reason to restrict attention to linear contracts in the first place." Arguing that nothing can be learned from a LEN model is akin to "throwing out the baby with the bath water." Although we present the general results first, we did use the LEN model as a guide to find them. The key was to tease out the information economics lurking behind the tempting explicit calculations of LEN models; we hope this approach will work in other cases of interest.

APPENDIX A

Proof of Proposition 1

The proof follows the lines of the proof of Proposition 18.6 in Christensen and Feltham (2005). We formulate the principal's problem of implementing a given action a by choosing an *ex*

[23] See also Liang and Nan (2008) for a related analysis in which the agent's action may affect the variance of a normally distributed performance measure.

ante optimal randomization strategy with probabilities $(\alpha, 1-\alpha)$ over two information systems (η^2, η^1) used to compensate the agent *ex post* with an optimal contract:

$$c(y|\eta) = \begin{cases} c_1(y) & \text{if } \eta = \eta^1 \\ c_2(y) & \text{if } \eta = \eta^2 \end{cases}. \tag{A1}$$

The principal optimizes over α and the contract. At contracting time, s/he announces the randomization parameter α (and commits to it) and offers the agent the contract in (A1). The principal's expected utility given randomization α over the two information systems and the contract is:

$$U^p(c,a|\alpha) = \alpha U^p(c_2,a|\eta^2) + (1-\alpha)U^p(c_1,a|\eta^1), \tag{A2}$$

$$U^p(c_i,a|\eta^i) = -\int c_i(y_i)\phi(y_i|a,\eta^i)dy_i. \tag{A3}$$

Knowing the randomization probability α over the two contracts, the agent's expected utility is:

$$U^a(c,a|\alpha) = \alpha U^a(c_2,a|\eta^2) + (1-\alpha)U^a(c_1,a|\eta^1), \tag{A4}$$

$$U^a(c_i,a|\eta^i) = \int u(c_i(y_i))k(a)\phi(y_i|a,\eta^i)dy_i - v(a). \tag{A5}$$

We assume that the first-order approach is valid, hence the agent's action choice is constrained by:

$$\nabla_a U^a(c,a|\alpha) = \alpha\nabla_a U^a(c_2,a|\eta^2) + (1-\alpha)\nabla_a U^a(c_1,a|\eta^1) = 0, \tag{A6}$$

$$\nabla_a U^a(c_i,a|\eta^i) = \int u(c_i(y_i))[k(a)\nabla_a\phi(y_i|a,\eta^i) + \nabla_a k(a)\phi(y_i|a,\eta^i)]dy_i - \nabla_a v(a). \tag{A7}$$

Let $\bar{U}$ denote the agent's lower bound on expected utility. Then, the principal's problem of inducing a given action a with an optimal contract and an optimal randomization strategy is characterized by the following Lagrangian:

$$\begin{aligned}\mathcal{L} = {} & \alpha U^p(c_2,a|\eta^2) + (1-\alpha)U^p(c_1,a|\eta^1) + \lambda[\alpha U^a(c_2,a|\eta^2) + (1-\alpha)U^a(c_1,a|\eta^1) - \bar{U}] \\ & + \mu\cdot[\alpha\nabla_a U^a(c_2,a|\eta^2) + (1-\alpha)\nabla_a U^a(c_1,a|\eta^1)].\end{aligned} \tag{A8}$$

The principal's optimization problem is $\max_{c_1,c_2,\alpha,\lambda,\mu}\mathcal{L}$, where the contracts $c_i(y_i)$ are in the set of functions bounded from below by $\underline{c}$. Optimizing the Lagrangian pointwise with respect to $c_1(y_1)$ and $c_2(y_2)$ for the common optimal multipliers λ and μ yields the following first-order conditions for the optimal compensation contracts, $i = 1,2$:

$$w(c_i(y_i)) = \begin{cases} \lambda k(a) + \mu\cdot[k(a)L(\phi(y_i),\eta^i|a) + \nabla_a k(a)] \\ \qquad \text{if } \lambda k(a) + \mu\cdot[k(a)L(\phi(y_i),\eta^i|a) + \nabla_a k(a)] \geq w(\underline{c}) \\ w(\underline{c}) \qquad \text{otherwise} \end{cases}, \tag{A9}$$

where $w(c) = 1/u'(c)$ is the agent's inverse marginal utility of consumption. For a risk-averse agent, the marginal utility of consumption $u'(c)$ is decreasing, so $w(c)$ is an increasing function.

The principal's expected utility and the agent's expected utilities are both linear functions of α so the optimal α is a corner solution in $[0,1]$; that is, $\alpha = 0$ or $\alpha = 1$. Differentiating the Lagrangian with respect to α yields:

$$B = U^p(c_2,a|\eta^2) + \lambda U^a(c_2,a|\eta^2) + \mu \cdot \nabla_a U^a(c_2,a|\eta^2) - [U^p(c_1,a|\eta^1) + \lambda U^a(c_1,a|\eta^1)$$
$$+ \mu \cdot \nabla_a U^a(c_1,a|\eta^1)] = \int [c_1(y_1) - u(c_1(y_1))m(y_1|a,\eta^1)]\phi(y_1|a,\eta^1)dy_1 - \int [c_2(y_2)$$
$$- u(c_2(y_2))m(y_2|a,\eta^2)]\phi(y_2|a,\eta^2)dy_2, \tag{A10}$$

where we substitute in the definitions of $U^p(c_i, a \mid \eta^j)$, $U^a(c_i, a \mid \eta^j)$, $\nabla_a U^a(c_1, a \mid \eta^1)$, we collect terms, and we denote:

$$m(y_i|a,\eta^j) = \lambda k(a) + \mu \cdot (k(a)L(\phi(y_i),\eta^j|a) + \nabla_a k(a)).$$

We now show that $B > 0$, so the principal optimally chooses $\alpha = 1$.

The optimal incentive contracts only depend on the information system η^j and the signal y_i through the likelihood ratio functions with the common multipliers λ and μ. Hence, we write both optimal contracts as $c_i(m) = cc(m)$:

$$cc(m) = \begin{cases} w^{-1}(m) & \text{if } m \geq w(\underline{c}) \\ \underline{c} & \text{otherwise} \end{cases}, \tag{A11}$$

where the likelihood measure m is defined as $m = \lambda k(a) + \mu \cdot [k(a)L(\phi(y_i), \eta^j \mid a) + \nabla_a k(a)]$ over its entire support. Only the distributions of the likelihood measure m depend on η^j, $\phi(m \mid a, \eta^2) \neq \phi(m \mid a, \eta^1)$, while the common function $cc(\cdot)$ includes the truncation due to the lower bound on compensation. We define $f(m) = cc(m) - u(cc(m))m$, so we can write B as:

$$B = \int f(m)\phi(m|a,\eta^1)dm - \int f(m)\phi(m|a,\eta^2)dm. \tag{A12}$$

The function $f(\cdot)$ is (strictly) concave (for $m > w(\underline{c})$). To see this, we differentiate:

$$f'(m) = cc'(m) - [u'(cc(m))cc'(m)m + u(cc(m))] = -u(cc(m)), \tag{A13}$$

where the second equality follows from $m = 1 / u'(cc(m))$ for $m \geq w(\underline{c})$, and $cc'(m) = 0$ for $m < w(\underline{c})$. Furthermore, note that $f''(m) = -u'(cc(m))cc'(m) < 0$ follows from $u' > 0$ and $cc'(m) > 0$ for $m > w(\underline{c})$. Finally, for likelihood measures $m' < w(\underline{c})$ such that $c(m') = \underline{c}$, we get that $f'(m') = -u(\underline{c}) > -u(cc(m))$ for $m > w(\underline{c})$, that is, for $cc(m) > \underline{c}$. Hence, the function $f(\cdot)$ is concave over the full support of the likelihood measure m, that is, over the interval $(-\infty, \infty)$, and strictly concave over the interval $[w(\underline{c}), \infty)$.

The final step is to note that if $\Phi(m \mid a, \eta^1)$ second-order stochastically dominates $\Phi(m \mid a, \eta^2)$ and the two distributions have the same mean, then for any concave function $f(\cdot)$ (be it increasing or decreasing), which is strictly concave over some range of nonzero probability at which the distributions do not coincide, it holds that, see Rothschild and Stiglitz (1970, Theorem 2):[24]

$$B = \int f(m)\phi(m|a,\eta^1)dm - \int f(m)\phi(m|a,\eta^2)dm > 0. \tag{A14}$$

□

The following lemma is useful in the proof of Proposition 2 below; see Billingsley (1995) for the probability concepts used in the lemma and its proof.

[24] Rothschild and Stiglitz (1970) only prove their Theorem 2 for distributions with compact support, but the result holds for general distributions under additional mild regularity conditions (Hadar and Russell 1969; Hanoch and Levy 1969).

Lemma A.1: Let $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^m$ be two vectors of independent zero-mean random variables defined on the product space $(\Omega,\mathcal{F},\mathcal{P})$ generated by the probability spaces $(\Omega_i, \mathcal{F}_i, \mathcal{P}_i)$ for the $\mathcal{F}_i$-measurable random variables x_i and y_i, $i = 1, \ldots, m$. If x_i second-order stochastically dominates y_i, for all $i = 1, \ldots, m$, then $a{\cdot}x$ second-order stochastically dominates $a{\cdot}y$ for any $a \in \mathbb{R}^m$.

Proof of Lemma A.1

Rothschild and Stiglitz (1970, Theorem 2) show that x_i SSD y_i if, and only if, there exists a $\mathcal{F}_i$-measurable random variable ε_i with $y_i \overset{d}{=} x_i + \varepsilon_i$, and $\mathrm{E}[\varepsilon_i \mid x_i] = 0$. The σ-fields $\mathcal{F}_i$ in the product space are independent; thus, if x_i SSD y_i, $i = 1, \ldots, m$, then there exists a $\mathcal{F}$-measurable vector of independent random variables $\varepsilon \in \mathbb{R}^m$ with $y_i \overset{d}{=} x_i + \varepsilon_i$ and $\mathrm{E}[\varepsilon_i \mid x_i] = 0$ for $i = 1, \ldots, m$. Because the random variables y_i and $x_i + \varepsilon_i$ have the same distribution functions, they also have the same characteristic functions defined by $\varphi_X(t) = \mathrm{E}[e^{itX}]$, $t \in \mathbb{R}$, i.e., $\varphi_{y_i}(t) = \varphi_{x_i+\varepsilon_i}(t)$. Since y is a vector of independent random variables, the characteristic function for $a{\cdot}y$ is $\varphi_{a\cdot y}(t) = \varphi_{y_1}(a_1 t) \ldots \varphi_{y_m}(a_m t)$. Similarly, $x + \varepsilon$ is a vector of independent random variables, so we have $\varphi_{a\cdot(x+\varepsilon)}(t) = \varphi_{x_1+\varepsilon_1}(a_1 t) \ldots \varphi_{x_m+\varepsilon_m}(a_m t)$. Since $\varphi_{y_i}(t) = \varphi_{x_i+\varepsilon_i}(t)$ implies that $\varphi_{y_i}(a_i t) = \varphi_{x_i+\varepsilon_i}(a_i t)$, for all $a_i \in \mathbb{R}$, we get that $\varphi_{a\cdot y}(t) = \varphi_{a\cdot(x+\varepsilon)}(t)$. But $a{\cdot}y$ and $a{\cdot}(x + \varepsilon)$ have the same characteristic functions, so they also have the same distribution functions (by the uniqueness theorem for characteristic functions, Billingsley 1995, Theorem 26.2): $a \cdot y \overset{d}{=} a \mid (x + \varepsilon) \overset{d}{=} a \cdot x + a \cdot \varepsilon$.

It remains to show that $\mathrm{E}[a \cdot \varepsilon | a \cdot x] = 0$. This is clearly satisfied if $\mathrm{E}[\varepsilon_i \mid a \cdot x] = 0$ for all i. By iterated expectations $\mathrm{E}[\varepsilon_i \mid a \cdot x] = \mathrm{E}[\mathrm{E}[\varepsilon_i \mid x] \mid a \cdot x]$. Since ε_i is $\mathcal{F}_i$-measurable and x_j is $\mathcal{F}_j$-measurable, ε_i and x_j are independent for $i \neq j$. Hence, $\mathrm{E}[\varepsilon_i \mid x] = \mathrm{E}[\varepsilon_i \mid x_i]$. The result now follows from the fact that $\mathrm{E}[\varepsilon_i \mid x_i] = 0$. Hence, there exists a $\mathcal{F}$-measurable random variable $z \equiv a{\cdot}\varepsilon$ such that $a \cdot y \overset{d}{=} a \cdot x + z$, where $\mathrm{E}[z \mid a \cdot x] = 0$. It then follows from Rothschild and Stiglitz (1970, Theorem 2) that $a{\cdot}x$ SSD $a{\cdot}y$ for any $a \in \mathbb{R}^m$. □

Proof of Proposition 2

Technological and stochastic independence imply that the zero-mean random variables $L_i(y_i \mid a_i, \eta) = \phi_{a_i}(y_i \mid a_i, \eta) / \phi(y_i \mid a_i, \eta)$ are independent random variables. If the task-specific performance measures are technologically and stochastically independent for both information systems, then the likelihood ratios for the two information systems are vectors of independent random variables defined on a common product space as in Lemma A.1. Hence, if $L_i(y_i \mid a_i, \eta^1)$ SSD $L_i(y_i \mid a_i, \eta^2)$ for all i, it follows from Lemma A.1 that $\mu \cdot L(y \mid a, \eta^1)$ SSD $\mu \cdot L(y \mid a, \eta^2)$ for any $\mu \in \mathbb{R}^m$. The result then follows from Proposition 1. □

Proof of Proposition 3

From the definition of likelihood ratios, it follows from (7a) that:

$$L(\phi(y_1), \eta^1 | a) = \int L(\phi(y_2), \eta^2 | a)\, \phi(y_2 | y_1, a) dy_2, \tag{A15}$$

where $\phi(y_2 \mid y_1, a) \equiv h(y_1, y_2)\phi(y_2 \mid a) / \phi(y_1 \mid a)$. Since (7a) also implies that $\int \phi(y_2 \mid y_1, a) dy_2 = 1$, $\phi(y_2 \mid y_1, a)$ is a conditional probability density function with respect to y_2 for any given y_1 and a. Thus, $L(\phi(y_1), \eta^1 \mid a) = \mathrm{E}[L(\phi, \eta^2 \mid a) \mid y_1, a]$. Define the random vector $\varepsilon \equiv L(\phi(y_2), \eta^2 \mid a) - \mathrm{E}[L(\phi, \eta^2 \mid a) \mid y_1, a]$. Then $L(\phi(y_2), \eta^2 \mid a) = L(\phi(y_1), \eta^1 \mid a) + \varepsilon$. Since this equality holds on all measurable sets, it also holds in distribution. Hence, for all $\mu \in \mathbb{R}^m$ and all $a \in A$ we have that $\mu \cdot L(\phi, \eta^2 \mid a) \overset{d}{=} \mu \cdot L(\phi, \eta^1 \mid a) + \mu \cdot \varepsilon$.

It remains to show that $E[\mu \cdot \varepsilon \mid \mu \cdot L(\phi, \eta^1 \mid a)] = 0$. It follows from iterated expectations that:

$$E[\mu \cdot \varepsilon | \mu \cdot L(\phi, \eta^1 | a)] = \mu \cdot E[E[\varepsilon | y_1, a] | \mu \cdot L(\phi, \eta^1 | a)].$$

The result now follows from the definition of ε:

$$E[\varepsilon | y_1, a] = E[L(\phi(y_2), \eta^2 | a) - E[L(\phi, \eta^2 | a) | y_1, a] | y_1, a] = 0.$$

Hence, there exists a random variable $z \equiv \mu \cdot \varepsilon$ such that $\mu \cdot L(\phi, \eta^2 \mid a) \overset{d}{=} \mu \cdot L(\phi, \eta^1 \mid a) + z$, where $E[z \mid \mu \cdot L(\phi, \eta^1 \mid a)] = 0$. It then follows from Rothschild and Stiglitz (1970, Theorem 2) that $\mu \cdot L(\phi, \eta^1 \mid a)$ SSD $\mu \cdot L(\phi, \eta^2 \mid a)$ for all μ and a. The result then follows from Proposition 1. □

Proof of Proposition 4

With square-root utility, the function $f(m) = cc(m) - u(cc(m))m$ is $f(m) = -m^2$, whenever the optimal compensation is interior. Because the optimal compensation is assumed to be interior over the full support of the likelihood measure, (A14) becomes a comparison of the variances of the corresponding likelihood measures. This proves the first result.

To determine the set of implementable actions, we start with the optimal contract (4) for additively separable square-root utility:

$$\sqrt{c(y)} = \lambda + \mu \cdot \frac{\nabla_a \phi}{\phi}(y | a, \eta). \tag{A16}$$

Multiplying by $\nabla_a \phi$ on both sides of (A16) and integrating, we have:

$$\int \sqrt{c(y)} \nabla_a \phi(y | a, \eta) dy = \int \left(\lambda + \mu \cdot \frac{\nabla_a \phi}{\phi}(y | a, \eta) \right) \frac{\nabla_a \phi(y | a, \eta)}{\phi(y | a, \eta)} \phi(y | a, \eta) dy. \tag{A17}$$

The right-hand side is $\Sigma L(\phi, \eta \mid a)\mu$ and the left- hand side is $\frac{1}{2}\nabla_a v(a)$ by the incentive compatibility constraint, thus $\Sigma L(\phi, \eta \mid a)\mu = \frac{1}{2}\nabla_a v(a)$. An action a is implementable if, and only if, $\nabla_a v(a)$ is in the image space of $\Sigma L(\phi, \eta | a)$. If $\Sigma L(\phi, \eta | a)$ is nonsingular, then $\mu = 1/2 \Sigma L(\phi, \eta \mid a)^{-1} \nabla_a v(a)$.

The square-root utility function also allows us to express the minimum expected compensation cost from (A16) as a function of the variance-covariance matrix of the likelihood ratio:

$$E[c(y) | a, \eta] = \lambda^2 + \Sigma L(\phi, \eta | a) \mu \cdot \mu. \tag{A18}$$

With square-root utility (assuming the optimal compensation is everywhere interior), we can solve for λ by exploiting the binding participation constraint. Integrating (A16) and using the binding participation constraint, we have:

$$\lambda = \int \sqrt{c(y)} \phi(y | a, \eta) dy = \tfrac{1}{2}[v(a) + \bar{U}]. \tag{A19}$$

Therefore, λ is a constant determined by the induced action (and the reservation utility). Substituting λ and μ into the expected compensation (A18), we have:

$$E[c(y) | a, \eta] = \frac{1}{4}[(v(a) + \bar{U})^2 + \Sigma L(\phi, \eta | a)^{-1} \nabla_a v(a) \cdot \nabla_a v(a)], \tag{A20}$$

which is a direct proof that the expected compensation cost of implementing an action a is lower whenever the variance-covariance matrix is larger. □

Proof of Proposition 5

Both distributions of the likelihood measure m for the two information systems η^1 and η^2 are normal, with mean $\lambda k(a) + \mu \cdot \nabla_a k(a)$ and variances $k^2(a)\Sigma L(\phi, \eta^1)\mu \cdot \mu$ and $k^2(a)\Sigma L(\phi, \eta^2)\mu \cdot \mu$, respectively. Hence, the distribution $\Phi(m \mid a, \eta^1)$ second-order stochastically dominates the distribution $\Phi(m \mid a, \eta^2)$ if, and only if, $\Sigma L(\phi, \eta^1)\mu \cdot \mu < \Sigma L(\phi, \eta^2)\mu \cdot \mu$, which is the case if $\Sigma L(\phi, \eta^2) - \Sigma L(\phi, \eta^1)$ is positive definite, that is, the VCM condition holds.

Following Proposition 3, if y_1 is a Blackwell garbling of y_2, the distribution function of $\mu \cdot L(\phi, \eta^2 \mid a)$ second-order stochastically dominates the distribution function of $\mu \cdot L(\phi, \eta^1 \mid a)$. For normally distributed y_1 and y_2, this means $\mu \cdot L(\phi, \eta^2 \mid a)$ has a higher variance than $\mu \cdot L(\phi, \eta^1 \mid a)$, that is, $\Sigma L(\phi, \eta^2)\mu \cdot \mu \geq \Sigma L(\phi, \eta^1)\mu \cdot \mu$, for all μ. Therefore, $\Sigma L(\phi, \eta^2) - \Sigma L(\phi, \eta^1)$ is a positive semi-definite matrix. This concludes the proof that $\Sigma L(\phi, \eta^2) \geq \Sigma L(\phi, \eta^1)$. □

Proof of Lemma 1

Writing explicitly the agent's certainty equivalent $\mathrm{CE}[c(y)|a,\eta]$, and because, with linear contracts, the action does not impact the variance of compensation, the individual rationality constraint, and the incentive compatibility constraint become:

$$\int c(y)\phi(y|a,\eta)dy - \tfrac{1}{2}r\int (c(y) - \mathrm{E}[c(y)|a,\eta])^2\phi(y|a,\eta)dy - \kappa(a) \geq \mathrm{CE}_0, \tag{A21}$$

$$\nabla_a \mathrm{CE}[c(y)|a,\eta] = \int c(y)\nabla_a\phi(y|a,\eta)dy - \nabla_a\kappa(a) = 0. \tag{A22}$$

The Lagrangian for the principal's problem is:

$$\mathcal{L}(c(y),\lambda,\mu) = -\mathrm{E}[c(y)|a,\eta] + \lambda(\mathrm{CE}[c(y)|a,\eta] - \mathrm{CE}_0) + \mu \cdot \nabla_a \mathrm{CE}(c(y)|a,\eta). \tag{A23}$$

The first-order condition with respect to $c(y)$ gives:[25]

$$-\phi(y|a,\eta) + \lambda[\phi(y|a,\eta) - r(c(y) - \mathrm{E}[c(y)|a,\eta])\phi(y|a,\eta)] + \mu \cdot \nabla_a\phi(y|a,\eta) = 0. \tag{A24}$$

Dividing through by $\phi(y{\cdot}a,\eta)$ gives the equivalent condition:

$$-1 + \lambda[1 - r(c(y) - \mathrm{E}[c(y)|a,\eta])] + \mu \cdot L(\phi(y),\eta|a) = 0. \tag{A25}$$

Integrating (A24) with respect to y gives $\lambda = 1$ because:

$$\int \nabla_a\phi(y|a,\eta)dy = 0 \text{ and } \int (c(y) - \mathrm{E}[c(y)|a,\eta])\phi(y|a,\eta)dy = 0. \tag{A26}$$

---

[25] Here we exploit again the fact that $c(y)$ is normally distributed. Because the vector space C, over which we optimize, only contains linear (affine) functions of y, an additional step is necessary to ensure that the first-order condition with respect to $c(y)$ is satisfied pointwise in y. Specifically, one calculates:

$$\partial\mathcal{L}/\partial c(y) = \mathrm{d}\mathcal{L}(c(y) + th(y))/\mathrm{d}t \text{ at } t = 0 \quad \text{for all } h(y) \in C.$$

This gives:

$$\int [-1 + \lambda(1 - r(c(y) - \mathrm{E}[c(y)|a,\eta])) + \mu \cdot \nabla_a\phi(y|a,\eta)/\phi(y|a,\eta)]h(y)\phi(y|a,\eta)dy = \int f(y)h(y)\phi(y|a,\eta)dy = 0,$$

where $f(y)$ denotes the term in square brackets. Note that $f(y)$ is the same as the left-hand side in (A25) and is a linear function of y because it is linear in $c(y)$ and in the likelihood ratio $L(\phi,\eta \mid a)$, which in turn is linear in y. Substituting $h(y) \equiv 1$ and $h(y) = f(y)$ in the integral form of the first-order condition, imply that $\mathrm{E}[f(y)] = \mathrm{var}(f(y)) = 0$, and since $f(y)$ is normally distributed, it follows that $f(y) \equiv 0$.

Substituting $\lambda = 1$ and solving for $c(y)$ in (A25) yields:

$$c(y) = \mathrm{E}[c(y)|a,\eta] + \frac{\mu}{r} \cdot L(\phi(y),\eta|a). \tag{A27}$$

We substitute the optimal linear contract (A27) into the agent's first-order incentive compatibility constraint (A22) to obtain:

$$\int \frac{\mu}{r} \cdot L(\phi(y),\eta|a) L(\phi(y),\eta|a)\phi(y|a,\eta)dy = \Sigma L(\phi,\eta)\frac{\mu}{r} = \nabla_a \kappa(a). \tag{A28}$$

This proves part (i) of the Lemma. There exists a μ that satisfies (A28) if, and only if, $\nabla_a \kappa(a)$ is in the image of the linear operator $\Sigma L(\phi,\eta)$, which is the same as the image of the linear operator M^T.[26] This proves part (ii) of the Lemma.

In what follows, assuming action a is implementable, we write for any μ that satisfies (A28):

$$\frac{\mu}{r} \in \Sigma L(\phi,\eta)^{-1} \nabla_a \kappa(a). \tag{A29}$$

When $\Sigma L(\phi,\eta)$ is singular and μ is not uniquely determined, $\Sigma L(\phi,\eta)^{-1}$ stands for the inverse image of the given vector, otherwise it denotes the inverse matrix. The participation constraint (A21) is binding in equilibrium, and we can use it to substitute the expected compensation in (A27) by the sum of the compensation risk premium and effort cost. Using (A27), the variance of the agent's compensation is:

$$\mathrm{var}(c(y)) = \mathrm{var}\left(\frac{\mu}{r} \cdot L(\phi,\eta|a)\right) = \Sigma L(\phi,\eta)\frac{\mu}{r} \cdot \frac{\mu}{r}. \tag{A30}$$

Substituting into (A27) gives the optimal linear contract that implements a:

$$c(y) = \mathrm{CE}_0 + \kappa(a) + \tfrac{1}{2} r \Sigma L(\phi,\eta)\frac{\mu}{r} \cdot \frac{\mu}{r} + \frac{\mu}{r} \cdot L(\phi(y),\eta|a). \tag{A31}$$

Thus, the principal's expected compensation cost for a given action is given by:

$$\mathrm{E}[c(y)|a,\eta] = \mathrm{CE}_0 + \kappa(a) + \tfrac{1}{2} r \Sigma L(\phi,\eta)\frac{\mu}{r} \cdot \frac{\mu}{r}. \tag{A32}$$

Substituting μ determined by (A29) in (A32), gives the optimal cost (16) and the optimal linear contract (15) for implementing action a. This proves part (iii) of the Lemma. To conclude, note that the optimal contract (15) and the expected cost (16) for inducing a are both uniquely determined for any μ that satisfies (A29), even when the multiplier μ is not uniquely determined, that is when a is implementable but $\Sigma L(\phi,\eta)$ is singular.[27] □

---

[26] To see this, note that $\Sigma L(\phi,\eta) = M^T \Sigma^{-1} M = A^T A$, where $A = \Sigma^{-1/2} M$. Because $\mathrm{Im}(A^T A) \subseteq \mathrm{Im}\, A^T$ and $\mathrm{rank}(A^T A) = \mathrm{rank}(A^T)$, it follows that $\mathrm{Im}(A^T A) = \mathrm{Im}(A^T)$. To conclude, note that $\mathrm{Im}(A^T) = \mathrm{Im}(M^T \Sigma^{-1/2}) = \mathrm{Im}(M^T)$.

[27] Indeed, let μ, μ' be two multipliers that both implement the same action a, that is both satisfy (A28). Then, $\Sigma L(\phi,\eta)(\mu - \mu') = 0$ implies $\Sigma L(\phi,\eta)\mu \cdot \mu = \Sigma L \mu' \cdot \mu'$ because:

$$\Sigma L(\phi,\eta)\mu \cdot \mu = \Sigma L(\phi,\eta)\mu' \cdot \mu = \mu' \cdot \Sigma L(\phi,\eta)\mu = \mu' \cdot \Sigma L(\phi,\eta)\mu' = \Sigma L(\phi,\eta)\mu' \cdot \mu'.$$

Furthermore, $L(\phi,\eta \mid a)$ is a zero mean normally distributed vector; this implies $(\mu - \mu') \cdot L(\phi,\eta \mid a)$ is normally distributed and has mean zero; finally, $\mathrm{var}((\mu - \mu') \cdot L(\phi,\eta \mid a)) = \Sigma L(\phi,\eta)(\mu - \mu') \cdot (\mu - \mu') = 0$ implies $(\mu - \mu') \cdot L(\phi,\eta \mid a) = 0$ because a normally distributed random variable with zero mean and variance is zero (almost) everywhere. By (A31), the optimal contract that induces a is the same for both μ and μ'.

Proof of Lemma 2

The proof is in two steps. In the first step, we show that, if $A \geq B \geq 0$ are two symmetric semi-positive definite matrices, then $\text{Im}(B) \subseteq \text{Im}(A)$.

We first show that $\text{Ker}(A) \subseteq \text{Ker}(B)$. For any vector $x \in \text{Ker}(A)$, we have $Ax = 0$, then $Ax \cdot x = 0$. The condition $A \geq B \geq 0$ implies $Ax \cdot x \geq Bx \cdot x \geq 0$, so we have $Bx \cdot x = 0$. Because B is semi-positive definite, there exists a square root $\sqrt{B}$. It follows that $0 = Bx \cdot x = \sqrt{B}x \cdot \sqrt{B}x$ and, thus $\sqrt{B}x = 0$. Hence, we have $\sqrt{B}\sqrt{B}x = Bx = 0$.

A vector x belongs to the kernel of a matrix if, and only if, x is orthogonal to the row space of the matrix. When the matrix is symmetric, x is also orthogonal to the column space of the matrix, which is the image of the associated linear operator. Therefore, because B is symmetric, $x \in \text{Im}(B)$, if, and only if, $x \perp \text{Ker}(B)$. Because we have established that $\text{Ker}(A) \subseteq \text{Ker}(B)$, then $x \perp \text{Ker}(B)$ implies $x \perp \text{Ker}(A)$ and $x \in \text{Im}(A)$. This completes the proof of the first step: if $x \in \text{Im}(B)$, then $x \in \text{Im}(A)$.

For the second step, recall that $A(\eta^i) = \{a \in \mathbb{R}^m \mid \nabla_a \kappa(a) \in \text{Im}(\Sigma L(\phi, \eta^i))\}$. If $a \in A(\eta^1)$ and $\Sigma L(\phi, \eta^2) \geq \Sigma L(\phi, \eta^1)$, then $\nabla_a \kappa(a) \in \text{Im}(\Sigma L(\phi, \eta^1)) \subseteq \text{Im}(\Sigma L(\phi, \eta^2))$. Thus, $a \in A(\eta^2)$, which concludes the proof. □

Proof of Proposition 6

For two information systems η^1, η^2, the actions that are implementable under both are characterized by:

$$\begin{aligned} A(\eta^1) \cap A(\eta^2) &= \{a \in \mathbb{R}^m | \nabla_a \kappa(a) \in \text{Im}(\Sigma L(\phi, \eta^1)) \cap \text{Im}(\Sigma L(\phi, \eta^2))\} \\ &= \{a \in \mathbb{R}^m | \nabla_a \kappa(a) \in \text{Im}(M_1^T) \cap \text{Im}(M_2^T)\}. \end{aligned}$$

In particular, if the row spaces of the sensitivity matrices M_1 and M_2 are equal, the actions that can be implemented with both information systems are the same. In this case, the restrictions of $\Sigma L(\phi, \eta^1)$ and $\Sigma L(\phi, \eta^2)$ to the subspace $\text{Im}(\Sigma L(\phi, \eta^1)) = \text{Im}(\Sigma L(\phi, \eta^2)) = \text{Im}(M_1^T) = \text{Im}(M_2^T)$ are both non-singular. We can then use the following lemma from Goldberger (1964, 38):

Lemma A.2: If A,B are positive definite matrices of the same order and $A \geq B$, then $B^{-1} \geq A^{-1}$.

Applying Lemma A.2 in comparing the expected compensation cost (16) of implementing the same action a proves the proposition. □

Proof of Lemma 3

Any implementable action a is determined by the multiplier μ from the optimal contract that implements a as follows (see (A28)):

$$a = \Sigma L(\phi, \eta)\frac{\mu}{r}. \tag{A33}$$

The principal's expected surplus from inducing a with an optimal contract is:

$$U^p = b \cdot a - \text{E}[c(y)|a, \eta] = b \cdot a - \text{CE}_0 - \tfrac{1}{2}a \cdot a - \tfrac{1}{2}r\Sigma L(\phi, \eta)\frac{\mu}{r} \cdot \frac{\mu}{r}. \tag{A34}$$

Recall that the principal's expected surplus when implementing a is well defined, even if $\Sigma L(\phi, \eta)$

is a singular matrix and μ is not uniquely determined by (A28).[28] Substituting (A33) in the principal's expected surplus gives:

$$U^p = b \cdot \Sigma L(\phi,\eta)\frac{\mu}{r} - \tfrac{1}{2}\Sigma L(\phi,\eta)\frac{\mu}{r} \cdot \Sigma L(\phi,\eta)\frac{\mu}{r} - \tfrac{1}{2}r\Sigma L(\phi,\eta)\frac{\mu}{r} \cdot \frac{\mu}{r} - \mathrm{CE}_0. \tag{A35}$$

Differentiating with respect to μ gives the following first-order condition with respect to μ:

$$\frac{1}{r}\Sigma L(\phi,\eta)b - \frac{1}{r}\Sigma L(\phi,\eta)^2\frac{\mu}{r} - \Sigma L(\phi,\eta)\frac{\mu}{r} = 0. \tag{A36}$$

Rearranging terms and solving for $\Sigma L(\phi,\eta)\mu/r$ gives:

$$a = \Sigma L(\phi,\eta)\frac{\mu}{r} = (\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b, \tag{A37}$$

where I denotes the $m \times m$ identity matrix. While μ may not be uniquely determined, $\mu = (\Sigma L(\phi, \eta) + rI)^{-1}b$ does satisfy the first-order condition, and the corresponding induced action is uniquely determined.[29] Finally, using (A37), we can explicitly calculate the principal's expected surplus. To complete the proof, it remains to show (18). We use (A37) to substitute into the principal's expected surplus:

$$\begin{aligned} U^p = {} & b \cdot (\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b - \tfrac{1}{2}(\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b \cdot (\Sigma L(\phi,\eta) \\ & + rI)^{-1}\Sigma L(\phi,\eta)b - \tfrac{1}{2}r(\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b \cdot \frac{\mu}{r} - \mathrm{CE}_0 \end{aligned} \tag{A38}$$

We next use the fact that $(\Sigma L(\phi, \eta) + rI)^{-1}$ and $\Sigma L(\phi,\eta)$ are symmetric and commute to further simplify the third term in (A38):

$$\begin{aligned} U^p = {} & (\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b \cdot \left[b - \tfrac{1}{2}(\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b\right] - \tfrac{1}{2}r(\Sigma L(\phi,\eta) \\ & + rI)^{-1}b \cdot (\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b - \mathrm{CE}_0. \end{aligned} \tag{A39}$$

Finally, factoring and simplifying gives the principal's expected surplus as:

$$\begin{aligned} U^p &= (\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b \cdot \left[b - \tfrac{1}{2}(\Sigma L(\phi,\eta) + rI)^{-1}(\Sigma L(\phi,\eta) + rI)b\right] - \mathrm{CE}_0 \\ &= \tfrac{1}{2}(\Sigma L(\phi,\eta) + rI)^{-1}\Sigma L(\phi,\eta)b \cdot b - \mathrm{CE}_0 = \tfrac{1}{2}(I - r(\Sigma L(\phi,\eta) + rI)^{-1})b \cdot b - \mathrm{CE}_0. \end{aligned} \tag{A40}$$

□

Proof of Proposition 8

To prove the proposition, it is sufficient to show: (i) ⇒ (ii) ⇒ (iii) ⇒ (iv) ⇒ (v) ⇒ (i). (i) ⇒ (ii). Let $z = y_2 - \mathrm{E}[y_2 \mid y_1, \hat{a}]$ for a given $\hat{a}$. The assumed parametrization of the performance measures implies that $\mathrm{E}[y_2 \mid y_1, \hat{a}] = M_2\hat{a} + \Sigma_{21}\Sigma_{11}^{-1}(y_1 - M_1\hat{a})$ and:

---

[28] The optimal contract is well defined by choosing $\mu / r = (\Sigma L(\phi, \eta) + rI)^{-1}b$, because any other multiplier μ' that satisfies (A33) has the property that $\Sigma L(\phi, \eta)(\mu - \mu') = 0$, which in turn implies $(\mu - \mu') \cdot L(\phi, \eta \mid a) = 0$. The matrix $\Sigma L(\phi,\eta) + rI$ is always non-singular because it is positive definite. The reason is that $\Sigma L(\phi,\eta)$ is non-negative definite, and rI is positive definite.

[29] This already exploits the particular structure of the Feltham and Xie (1994) model, where the agent has preferences over how effort is allocated among tasks. While in general a may not be uniquely determined by the first-order condition $\Sigma L(\phi, \eta)\mu / r = \nabla_a\kappa(a)$, this is not a problem when $\kappa(a) = \frac{1}{2}a \cdot a$ because $\nabla_a\kappa(a) = a$ in this case. The specification $\kappa(a) = \frac{1}{2}a \cdot a$ replaces, through a simple linear re-scaling, without loss of generality, any quadratic effort cost of the form $\kappa(a) = Ka \cdot a$ where K is positive-definite.

$$z(\hat{a}) = (M_2 - \Sigma_{21}\Sigma_{11}^{-1}M_1)(a - \hat{a}) + \varepsilon_2 - \Sigma_{21}\Sigma_{11}^{-1}\varepsilon_1. \quad \text{(A41)}$$

We start with the observation that y_2 has no incremental value over y_1 if, and only if, $z(\hat{a})$ has no incremental value over y_1. This is true because any contract that can be written on (y_1, y_2) can be equivalently written on $(z(\hat{a}), y_1)$ and *vice versa*. By construction, z is independent of y_1, thus $\phi(y_1, z \mid a, \hat{a}) = \phi(y_1 \mid a)\phi(z \mid a, \hat{a})$ and the likelihood ratio is additively separable:

$$L(\phi|y_1,z(\hat{a}),a) = L(\phi|y_1,a) + L(\phi|z(\hat{a}),a). \quad \text{(A42)}$$

The likelihood ratios are linear in the independent normally distributed random variables y_1 and $z(\hat{a})$, and are independent normally distributed random variables as well with additively separable variance-covariance matrices:

$$\Sigma L(\phi|y_1,z(\hat{a})) = \Sigma L(\phi|y_1) + \Sigma L(\phi|z(\hat{a})). \quad \text{(A43)}$$

Our assumption that y_2 has no incremental value over y_1 implies that $z(\hat{a})$ has no incremental value over y_1. Thus, $\Sigma L(\phi \mid z(\hat{a})) = 0$; otherwise $\Sigma L(\phi \mid y_1, z(\hat{a})) > \Sigma L(\phi \mid y_1)$ implies (by Corollary 5) that $z(\hat{a})$ has incremental value over y_1 and contradicts our assumption. Furthermore, $\Sigma L(\phi \mid z(\hat{a})) = 0$ implies that $L(\phi \mid z(\hat{a}), a) = 0$ because it is a normally distributed variable with mean and variance-covariance equal to zero. This is only possible if $\nabla_a \phi(z(\hat{a}) \mid a) = 0$ for all a, which implies $M_2 - \Sigma_{21}\Sigma_{11}^{-1}M_1 = 0$ and proves the desired result with $Q = \Sigma_{21}\Sigma_{11}^{-1}$.

(ii) $\Rightarrow$ (iii) From $\mathrm{E}[y_2 \mid y_1, a] = M_2 a + \Sigma_{21}\Sigma_{11}^{-1}(y_1 - M_1 a)$ and the assumption that $M_2 = QM_1$, $\Sigma_{21} = Q\Sigma_{11}$, we have $\mathrm{E}[y_2 \mid y_1, a] = \Sigma_{21}\Sigma_{11}^{-1}y_1$. Thus, we can write:

$$y_2 = \mathrm{E}[y_2|y_1,a] + y_2 - \mathrm{E}[y_2|y_1,a] = \Sigma_{21}\Sigma_{11}^{-1}y_1 + \varepsilon$$

where $\varepsilon = \varepsilon_2 - \Sigma_{21}\Sigma_{11}^{-1}\varepsilon_1$ is independent of y_1.

(iii) $\Rightarrow$ (iv) It is sufficient to show that the conditional density $\phi(y_2 \mid y_1, a)$ does not depend on a because $\phi(y_1, y_2 \mid a) = \phi(y_2 \mid y_1, a)\phi(y_1 \mid a)$; because the signals are normally distributed and a does not affect their variance, it is sufficient to prove that $\mathrm{E}[y_2 \mid y_1, a]$ does not depend on a. But $\mathrm{E}[y_2 \mid y_1, a] = Qy_1 + \mathrm{E}[\varepsilon \mid y_1, a] = Qy_1$ only depends on y_1 and this concludes the argument.

(iv) $\Rightarrow$ (v) A direct calculation shows that $L(\phi \mid y_1, y_2, a) = L(\phi \mid y_1, a)$ for all a; it follows that $\Sigma L(\phi \mid y_1, y_2) = \Sigma L(\phi \mid y_1)$.

(v) $\Rightarrow$ (i) From Corollary 5, the principal weakly prefers (y_1, y_2) to y_1 because $\Sigma L(\phi \mid y_1, y_2) \geq \Sigma L(\phi \mid y_1)$ and weakly prefers y_1 to (y_1, y_2) because $\Sigma L(\phi \mid y_1, y_2) \leq \Sigma L(\phi \mid y_1)$. Thus, the principal is indifferent between the two information systems and y_2 has no incremental value over y_1.

To conclude, the sufficient statistic condition (iv) implies that $h(y_1, y_2) = \phi(y_1, y_2 \mid a) / \phi(y_1 \mid a)$ satisfies the Blackwell garbling conditions in (7). □

APPENDIX B

In this appendix we examine a multi-task setting in which the agent's action choice is discrete and, thus, the incentive constraints cannot be substituted with a first-order condition. Our approach in characterizing the optimal contracts follows that of Grossman and Hart (1983). We show that the result in Proposition 1 still applies with a slightly different definition of the likelihood ratios. Specifically, let $A = \{a_j \in \mathbb{R}^m\}$ be a finite set of (multi-task) actions available to the agent. As in the proof of Proposition 1, we examine the principal's problem of implementing a given action $a \in A$ with an optimal contract by choosing an *ex ante* optimal randomization strategy with probabilities $(\alpha, 1 - \alpha)$ over two information systems (η^2, η^1) used to compensate the agent *ex post* with

contracts c_2 and c_1, respectively. Knowing the randomization probability α over the two contracts, the agent's expected utility is:

$$U^a(c,a|\alpha) = \alpha U^a(c_2,a|\eta^2) + (1-\alpha)U^a(c_1,a|\eta^1), \tag{B1}$$

$$U^a(c_i,a|\eta^i) = \int u(c_i(y_i))k(a)\phi(y_i|a,\eta^i)dy_i - v(a). \tag{B2}$$

The agent's incentive compatibility constraints are:

$$U^a(c,a|\alpha) \geq U^a(c,a_j|\alpha), \quad \text{for all } a_j \in A, \tag{B3}$$

with associated non-negative Lagrange-multipliers μ_j.

The principal's problem of inducing action a with an optimal contract and an optimal randomization strategy is characterized by the following Lagrangian:

$$\begin{aligned}\mathcal{L} = {} & \alpha U^p(c_2,a|\eta^2) + (1-\alpha)U^p(c_1,a|\eta^1) + \lambda[\alpha U^a(c_2,a|\eta^2) + (1-\alpha)U^a(c_1,a|\eta^1) - \bar{U}] \\ & + \sum\nolimits_{a_j \in A} \mu_j[\alpha[U^a(c_2,a|\eta^2) - U^a(c_2,a_j|\eta^2)] + (1-\alpha)[U^a(c_1,a|\eta^1) - U^a(c_1,a_j|\eta^1)]].\end{aligned} \tag{B4}$$

As in the proof of Proposition 1, optimizing the Lagrangian pointwise with respect to $c_1(y_1)$ and $c_2(y_2)$ (for the common optimal multipliers λ and μ_j) yields the following first-order conditions for the optimal compensation contracts for $i = 1,2$:

$$w(c_i(y_i)) = \begin{cases} m(y_i|a,\eta^i) & \text{if } m(y_i|a,\eta^i) \geq w(\underline{c}) \\ w(\underline{c}) & \text{otherwise} \end{cases}, \tag{B5}$$

where we define:

$$m(y_i|a,\eta^i) = \lambda k(a) + \sum_{a_j \in A} \mu_j[k(a_j)L(\phi(y_i),\eta^i|a,a_j) + k(a) - k(a_j)], \tag{B6}$$

$$L(\phi(y_i),\eta^i|a,a_j) = \frac{\phi(y_i|a,\eta^i) - \phi(y_i|a_j,\eta^i)}{\phi(y_i|a,\eta^i)}. \tag{B7}$$

Differentiating the Lagrangian with respect to α yields:

$$\begin{aligned}B = {} & U^p(c_2,a|\eta^2) + \lambda U^a(c_2,a|\eta^2) + \sum\nolimits_{a_j \in A} \mu_j[U^a(c_2,a|\eta^2) - U^a(c_2,a_j|\eta^2)] - \Big[U^p(c_1,a|\eta^1) \\ & + \lambda U^a(c_1,a|\eta^1) + \sum\nolimits_{a_j \in A} \mu_j[U^a(c_1,a|\eta^1) - U^a(c_1,a_j|\eta^1)]\Big] = \int [c_1(y_1) \\ & - u(c_1(y_1))m(y_1|a,\eta^1)]\phi(y_1|a,\eta^1)dy_1 - \int [c_2(y_2) - u(c_2(y_2))m(y_2|a,\eta^2)]\phi(y_2|a,\eta^2)dy_2,\end{aligned}$$

where for the second equality, we substitute in the definitions of $U^p(c_i, a \mid \eta^i)$, $U^a(c_i, a \mid \eta^i)$ and collect terms.

Denote $\mu(a) = (\mu_j k(a_j))_{a_j \in A}$, $L(\phi(y_i), \eta^i \mid a) = (L(\phi(y_i), \eta^i \mid a, a_j))_{a_j \in A}$, and $\Delta k(a) = (k(a) - k(a_j))_{a_j \in A}$. Then, the likelihood measure is $m = \lambda k(a) + \mu(a) \cdot (L(\phi(y_i), \eta^i \mid a) + \Delta k(a))$ and:

$$c_i(m) = \begin{cases} w^{-1}(m) & \text{if } m \geq w(\underline{c}) \\ \underline{c} & \text{otherwise} \end{cases} . \tag{B8}$$

With these redefinitions of the likelihood ratio and the likelihood measure, we follow the same steps as in the proof of Proposition 1 and we obtain its discrete equivalent as follows.

Proposition B.1: Assume the agent's utility is either additively or multiplicatively separable in compensation and effort. If a is an implementable action under two information systems η^1 and η^2, then there exists a vector $\mu(a)$ such that, if the distribution function of $\mu(a) \cdot L(\phi, \eta^1 \mid a)$ second-order stochastically dominates the distribution of $\mu(a) \cdot L(\phi, \eta^2 \mid a)$, then the principal prefers η^2 to η^1 for implementing action a.

This discrete equivalent of Proposition 1 demonstrates that the key assumption in ranking information systems in a multi-task agency is the second-order stochastic dominance of $\mu(a) \cdot L$ for the likelihood vector L defined in (B7). In general, this likelihood vector is of the same dimension as the number of actions available in the discrete set A. If we eliminate the non-binding incentive compatibility constraints (for which $\mu_j = 0$), we can restrict our attention to a likelihood vector of dimension equal to the number of binding incentive constraints.

To obtain something even closer to our Proposition 1, we assume that only the incentive compatibility constraints for the actions "adjacent" to the desired one are binding. To define the actions adjacent to the action $a = (a_1, \ldots, a_m) \in \mathbb{R}^m$ that the principal wants to implement, suppose that the action set A is rich enough so that $\hat{a}_k = (a_1, \ldots, a_{k-1}, a_k^-, a_{k+1}, \ldots, a_m) \in A$ for all $k = 1, \ldots, m$; that is, it contains all the actions obtained by substituting the effort level on task k, a_k, by the next less costly one a_k^-, while maintaining the effort levels on all the other tasks. The actions $\hat{a}_k$ are the actions adjacent to a in the m different "directions" determined by the m tasks. If we assume that all the incentive compatibility constraints for $a_j \in A \setminus \{\hat{a}_1, \ldots, \hat{a}_m\}$ are not binding, then the multiplier $\mu(a)$ in Proposition B.1 is an m-dimensional vector as in Proposition 1.

It takes a number of strong assumptions to ensure that only the adjacent incentive constraints are binding even in single-task settings (for example, spanning or convexity of distribution function conditions, such as in Grossman and Hart 1983). In single-task settings, the conditions sufficient for the adjacent incentive constraints to be binding are very similar to those that are sufficient for the first-order approach to apply (Christensen and Feltham 2005, Appendix 17A). It is beyond the scope of this study to derive such conditions in multi-task settings. However, with only two possible actions (such as high and low), there is only one incentive constraint and that is binding by assumption.

REFERENCES

Adams, C. P. 2006. Optimal team incentives with CES production. *Economics Letters* 92 (1): 143–148.

Amershi, A. H., R. D. Banker, and S. M. Datar. 1990. Economic sufficiency and statistical sufficiency in the aggregation of accounting signals. *The Accounting Review* 65 (1): 113–130.

Antle, R., and J. S. Demski. 1988. The controllability principle in responsibility accounting. *The Accounting Review* 63 (4): 700–718.

Banker, R. D., and S. M. Datar. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* (1): 21–39.

Billingsley, P. 1995. *Probability and Measure*. 3rd edition. New York, NY: John Wiley & Sons, Inc.

Bushman, R. M., and R. J. Indjejikian. 1993. Accounting income, stock price, and managerial compensation. *Journal of Accounting and Economics* 16 (1–3): 3–23.

Christensen, P. O., and G. A. Feltham. 2005. *Economics of Accounting: Volume II—Performance Evaluation*. Springer Series in Accounting Scholarship. New York, NY: Springer Science + Business Media Inc.

———, ———, and F. Şabac. 2003. Dynamic incentives and responsibility accounting: A comment. *Journal of Accounting and Economics* 35 (3): 423–436.

———, ———, and F. Şabac. 2005. A contracting perspective on earnings quality. *Journal of Accounting and Economics* 39 (2): 265–294.

Conlon, J. R. 2009. Two new conditions supporting the first-order approach to multisignal principal-agent problems. *Econometrica* 77 (1): 249–278.

Datar, S. M., S. C. Kulp, and R. A. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39 (1): 75–92.

Demski, J. S., J. C. Fellingham, and H. H. Lin. 2007. Efficient evaluation. Working paper, University of Florida.

Dewatripont, M., I. Jewitt, and J. Tirole. 1999. The economics of career concerns, Part I: Comparing information structures. *The Review of Economic Studies* 66 (1): 183–198.

Dutta, S., and S. Reichelstein. 1999. Asset valuation and performance measurement in a dynamic agency setting. *Review of Accounting Studies* 4 (3–4): 235–258.

———, and ———. 2002. Controlling investment decisions: Depreciation and capital charges. *Review of Accounting Studies* 7 (2–3): 253–281.

———, and ———. 2003. Leading indicator variables, performance measurement and long-term versus short-term contracts. *Journal of Accounting Research* 41 (5): 837–866.

Feltham, G. A., and M. G. Wu. 2000. Public reports, information acquisition by investors, and management incentives. *Review of Accounting Studies* 5 (2): 155–190.

———, and ———. 2001. Incentive efficiency of stock versus options. *Review of Accounting Studies* 6 (1): 1–38.

———, and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69 (3): 429–453.

Gibbons, R., and K. J. Murphy. 1992. Optimal incentive contracts in the presence of career concerns: Theory and evidence. *The Journal of Political Economy* 100 (3): 468–504.

Goldberger, A. S. 1964. *Econometric Theory*. New York, NY: John Wiley & Sons, Inc.

Grossman, S. J., and O. D. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* 51 (1): 7–45.

Hadar, J., and W. R. Russell. 1969. Rules for ordering uncertain prospects. *The American Economic Review* 59 (1): 25–34.

Hanoch, G., and H. Levy. 1969. Efficiency analysis of choices involving risks. *The Review of Economic Studies* 36 (107): 335–346.

Hellwig, M. F., and K. M. Schmidt. 2002. Discrete-time approximations of the Holmström-Milgrom Brownian motion model of intertemporal incentive provision. *Econometrica* 76 (6): 2225–2264.

Hemmer, T. 2004. Lessons lost in linearity: A critical assessment of the general usefulness of LEN models in compensation research. *Journal of Management Accounting Research* 16 (1): 149–162.

Holmström, B. 1979. Moral hazard and observability. *The Bell Journal of Economics* 10 (1): 74–91.

———, and P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law Economics and Organization* 7 (Special issue): 24–52.

———, and ———. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55 (2): 303–328.

Huddart, S., and P. J. Liang. 2005. Profit sharing and monitoring in partnerships. *Journal of Accounting and Economics* 40 (1–3): 153–187.

Indjejikian, R., and D. Nanda. 1999. Dynamic incentives and responsibility accounting. *Journal of Accounting and Economics* 27 (2): 177–201.

Jewitt, I. 1988. Justifying the first-order approach to principal-agent problems. *Econometrica* 56 (5): 1177–1190.

Kim, O., and Y. Suh. 1993. Incentive efficiency of compensation based on accounting and market performance. *Journal of Accounting and Economics* 16 (1–3): 25–53.

Kim, S. K. 1995. Efficiency of an information system in an agency model. *Econometrica* 63 (1): 89–102.

———, and Y. S. Suh. 1991. Ranking of accounting information systems for management control. *Journal of Accounting Research* 29 (2): 386–396.

Lambert, R. A. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (1–3): 3–87.

Liang, P. J., and L. Nan. 2008. Limited managerial attention and endogenous precision of performance measures. Working paper, Carnegie Mellon University.

——, M. Rajan, and K. Ray. 2008. Optimal team size and monitoring in organizations. *The Accounting Review* 83 (3): 789–822.

Marschak, J., and K. Miyasawa. 1968. Economic comparability of information systems. *International Economic Review* 9 (2): 137–174.

McLaughlin, K. J. 1994. Individual compensation and firm performance: The economics of team incentives. Working paper, The University of Chicago.

Meyer, M. A., and J. Vickers. 1997. Performance comparisons and dynamic incentives. *The Journal of Political Economy* 105 (3): 547–581.

Mirrlees, J. A. 1999. The theory of moral hazard and unobservable behaviour: Part I. *The Review of Economic Studies* 66 (1): 3–21. *This paper was completed in October 1975 but never previously published (eds.).*

Rogerson, W. P. 1985. The first-order approach to principal-agent problems. *Econometrica* 53 (6): 1357–1368.

Rothschild, M., and J. E. Stiglitz. 1970. Increasing risk: I. A definition. *Journal of Economic Theory* 2 (3): 225–243.

Şabac, F. 2007. Dynamic agency with renegotiation and managerial tenure. *Management Science* 53 (5): 849–864.

—— 2008. Dynamic incentives and retirement. *Journal of Accounting and Economics* 46 (1): 172–200.

Sinclair-Desgagné, B. 1994. The first-order approach to multi-signal principal-agent problems. *Econometrica* 62 (2): 459–465.

—— 2009. Ancillary statistics in principal-agent models. *Econometrica* (1): 279–281.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1577–1615

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1577

The Persistence and Market Valuation of Recurring Nonrecurring Items

William Cready
The University of Texas at Dallas

Thomas J. Lopez
The University of Alabama

Craig A. Sisneros
University of Colorado Denver

ABSTRACT: This study focuses on the persistence and market value implications of a subset of nonrecurring charges that are atypical due to repeated occurrence. The increased recurrence of supposedly nonrecurring items perhaps reflects managerial shifting of (more permanent) ordinary expenses to a transitory category or, alternatively, may reflect an environment where these items naturally occur more frequently. Either scenario suggests that these repetitive charges have future earnings implications dramatically different from truly nonrecurring events and should therefore be valued more like a recurring component of earnings. Consistent with this notion, we find that as the frequency of reporting negative special items increases (measured by the presence of multiple prior charges), the persistence of these items significantly increases with respect to future earnings. Our evidence also suggests that the valuation multiple on such charges increases with frequency. That is, the market values "recurring nonrecurring" items more like the other components of recurring earnings.

Keywords: *nonrecurring items; special items; restructuring; frequency; persistence; market value.*

Data Availability: *Data are available from public sources indicated in the text.*

We thank Larry Brown, Dennis Chambers, Ted Christensen, Don Hermann, Scott Jackson, Peter Johnson, Denise Leggett, Sarah Rector, Lynn Rees, Austin Reitenga, K. Sivaramakrishnan, Kay Stice, Mary Stone, and Charlotte Wright for their valuable comments. We also thank workshop participants at The University of Alabama, Arizona State University, Georgia State University, University of Kentucky, Oklahoma State University, University of South Carolina, Texas A&M University, and the 2005 FARS Midyear Meeting (especially the discussant, Katherine Schipper). We have particularly benefited from the helpful comments of two anonymous reviewers and Steve Kachelmeier (editor).

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: December 2007
Accepted: November 2009
Published Online: August 2010

I. INTRODUCTION

Burgstahler et al. (2002) provide evidence consistent with the conclusion that negative special items are largely transitory. The empirical estimates from their earnings prediction model confirm that negative special items (NSIs) are more transitory than the remaining components of earnings. Consistent with this notion, a substantial body of empirical research finds that the market places a substantially lower valuation multiple (response coefficient) on accounting income components *classified* as nonrecurring (e.g., Strong and Walker 1993; Elliott and Hanna 1996). This lower valuation is consistent with the notion that investors view these items as finite-horizon events and supports the accounting reporting policy that allows material nonrecurring income components to be identified separately in the income statement. We extend this literature by investigating the persistence and market valuation of a subset of "nonrecurring" items that are atypical due to repeated occurrence.[1]

The financial press is replete with articles addressing the increased incidence of repetitive "nonrecurring" charges. For example, according to one Merrill Lynch analyst, "special items are so prevalent now that they're not special anymore" (Fowler 2006, A1). DiMartino (2002, D1) writes, "Corporate America has been reporting so many *nonrecurring* restructuring charges that it's tough to see where routine operations end and restructuring begins" (emphasis added). Barton Biggs, former chief global strategist for Morgan Stanley, stated, "One, maybe even two 'nonrecurring charges' are understandable and possibly bullish for a company's stock price, because they clear the decks for future earnings gains, but how can repeated write-offs be nonrecurring or extraordinary? How can investors in the future believe that reported earnings are for real and won't be canceled by subsequent write-offs?" (Smith and Lipin 1996, A1). Hence, it is of particular interest to improve our understanding of the persistence characteristics of these repetitive "nonrecurring" charges and, more importantly, how the market values such charges.[2]

Empirically, we show that the frequency of prior reported NSIs is significantly associated with the future reporting of NSIs. We find firms that report no NSIs over the prior 12 quarters report significantly fewer NSIs over the subsequent 12 quarters than firms that report more than four prior charges. We also estimate earnings persistence models that relate current earnings and NSIs to earnings in the first and fourth subsequent quarters. Our evidence suggests that as the frequency of prior reported NSIs increases, the persistence of these items with respect to future earnings also increases. Overall, our evidence suggests that the frequency of prior NSIs has persistence implications for future earnings and NSIs.

Consistent with Elliott and Hanna (1996), we find in a regression of market returns on unexpected earnings and NSIs that the coefficient on unexpected earnings declines as the frequency of prior NSIs increases. More importantly for our analysis, we find that the coefficient on NSIs *increases* as the frequency of prior reported special items increases. This result, contrary to the findings of Elliott and Hanna (1996), suggests that the market values recurring NSIs as more

---

[1] This study focuses on the implications of negative special items (NSIs) rather than all special items for three reasons. First, most special items are income-decreasing. Burgstahler et al. (2002) report more than twice as many NSIs in their sample compared to positive special items. Second, our examination directly extends Elliott and Hanna (1996), which focuses solely on NSIs. Third, in our own analyses of positive special items, we find no evidence of empirical regularities meriting further investigation.

[2] In addition, standard-setters have also expressed concern with the clarity of financial statements with respect to the reporting of infrequent or unusual items (nonrecurring items). Reporting of these items has generally been governed by a relatively old standard, Accounting Principles Board (APB) Opinion No. 30, *Reporting the Results of Operations* (APB 1973). While it appears standard-setters are not considering new guidance on how these items are reported, there is substantial interest in clarifying the implications of these items. For example, in a joint project of the FASB and IASB on financial statement presentation, the preliminary views document states, "An entity should present as part of that schedule [a schedule that reconciles the statement of cash flows to the statement of comprehensive income] information about amounts related to an unusual or infrequent event or transaction" (FASB and IASB 2008).

persistent (less transitory) than truly nonrecurring charges. Although the coefficient on NSIs increases with prior reporting frequency, our evidence also suggests that the coefficient on unexpected earnings uniformly exceeds the special items coefficients in all cases. That is, even if a special item is recurring in nature, the market still values these items as less persistent than the other components of earnings.

We further extend this analysis by focusing on a specific type of NSI—restructuring charges. Restructuring charges are of particular interest for several reasons. First, restructuring charges commonly reflect a specific type of decision that is not only of established interest within the academic accounting and economics literatures (e.g., Dechow et al. 1994; Dial and Murphy 1995; Chaney et al. 1999; Adut et al. 2003; Clement et al. 2007), but are also of broad interest in the business media (e.g., Bleakley 1995; DiMartino 2002; Smith and Lipin 1996; Stewart 2002). Second, because restructurings focus on firm downsizing, these events constitute a comparatively homogeneous subsample within the highly heterogeneous population of special items.

Consistent with the results on special items, we find that the market attaches less weight to restructuring charges than to earnings before restructuring charges, the coefficient on unexpected earnings declines as the frequency of prior restructuring charges increases, and restructuring charge coefficients increase monotonically with the number of prior reported restructuring charges. In fact, estimated restructuring charge coefficients beyond the second in sequence become indistinguishable from the coefficients on earnings. Overall, our evidence suggests the market values repetitive restructuring charges more like permanent components of earnings than transitory charges.

Our evidence is broadly consistent with the idea that managers find it attractive to classify ordinary recurring operating expenses as "special items." The prior literature (e.g., Lipe 1986; Elliott and Hanna 1996; Francis et al. 1996; Bradshaw and Sloan 2002; McVay 2006) suggests that the further a line item is from sales on the income statement, the less weight the market attaches to the individual line item. Thus, the stockholders of the firm potentially benefit from the lower valuation multiple attached to NSIs (which are generally reported further from sales than operating expenses). Our evidence also suggests that "nonrecurring" items have become more persistent for some firms. More importantly, the market appears responsive to this phenomenon, in that such "recurring nonrecurring" charges are valued more like ordinary components of income. In addition, our results clearly show that, as firms increasingly report NSIs (restructuring charges), the persistence of adjusted earnings decreases and the market responds to this lower persistence as evidenced by the decreasing valuation on earnings.

Section II discusses the prior related literature. Section III develops the research design. Section IV presents the sample selection and descriptive statistics. Section V presents empirical results, and Section VI briefly summarizes and concludes.

II. PRIOR RELATED LITERATURE

The business media and analysts typically treat restructuring charges as nonrecurring costs, commonly referring to "earnings before restructuring charges" as the primary measure of earnings.[3] Bricker et al. (1995, 547) report that analysts focus on the "*core* operating results of companies" (emphasis in original) and that such analysis "typically involve(s) the removal of items from continuing operations that analysts judge are nonrecurring" (emphasis in the original). Lalli and Lim (1997, 76) state, "these so-called special, nonrecurring losses are generally ignored

[3] For example, the *New York Times* on September 6, 2002 (Section C; Page 4, emphasis added) reported, "The Procter & Gamble Company said yesterday that it expected *earnings before restructuring charges* to increase by about 15 percent in the current quarter ... That would beat Wall Street expectations. Analysts surveyed by Thomson First Call expected *core earnings* per share of $1.09, up 13.5 percent from 96 cents a share a year ago."

by both Wall Street analysts and shareholders, they are viewed as isolated one-time events (even when companies seem to claim them every two to three years)."

Strong and Walker (1993) examine the association between annual stock returns and reported earnings for United Kingdom industrial companies. They report that the earnings response coefficient on "pre-exceptional" earnings is significantly greater than the coefficient on "exceptional" earnings, where exceptional earnings are defined as material unusual or infrequent items. Similarly, Parkash (1995) demonstrates that permanent unexpected earnings exhibit a stronger relation with stock returns than temporary unexpected earnings. Elliott and Hanna (1996) find that large NSIs induce far smaller announcement period price responses than do earnings before NSIs. These results are consistent with Kormendi and Lipe (1987) and Collins and Kothari (1989), who conclude the market response to earnings is positively related to earnings persistence.

The notion that special items, which include restructuring charges, are transitory is central to policy perspectives on how such items should be reported. However, consistent with anecdotal reports (Bleakley 1995; Stewart 2002; DiMartino 2002; Fowler 2006), empirical research provides evidence that (1) the incidence of repeat "nonrecurring" charges is pervasive, and (2) such charges possibly include ordinary income components. The repetitive nature of these charges is evident in at least three prior studies. Of the 91 firms in Dechow et al. (1994), 45 report multiple charges over the eight-year sample time period, 26 of which report three or more restructuring charges. Similarly, 22 percent of the large write-off firms (i.e., negative special item firms) in Elliott and Hanna (1996) report three or more such charges over the prior 12 quarters. Approximately two-thirds of the 149 firms in Adut et al. (2003) report a temporally proximate restructuring charge (a second charge within two years). Dechow et al. (1994, 153) report the degree to which compensation is shielded from the restructuring charge decreases with the frequency of such charges. They interpret this finding as consistent with the notion that such charges "are more likely to contain ongoing operating costs." Adut et al. (2003) refine this aspect of the Dechow et al. (1994) analysis and find that temporally proximate restructuring charges do not result in shielding. That is, repetitive restructuring charges are treated like other operating expenses for executive compensation purposes.

III. RESEARCH DESIGN

Overview

In a market-based study, Elliott and Hanna (1996, 136) state, "firms might transfer normal components of operating expenses into the special item" as an underlying rationale for why such items might adversely impact the market valuation of income before special items. While this makes sense for earnings, the suggestion that repetitive nonrecurring charges are likely to contain regular components of operating expenses leads to the following questions related to special items: If normal components of operating expenses are shifted into special items, then does the market recognize this and attach a higher valuation to these costs? Similarly, if seemingly nonrecurring items become recurring in nature, then does the market recognize and value this phenomenon accordingly? Conversely, is the market either ignoring or somehow fooled by the fact that these repetitive costs are identified in the financial statements as nonrecurring items? The existing empirical evidence suggests the latter may be the case. Specifically, Elliott and Hanna (1996) find that the response coefficient on NSIs becomes insignificant as these items become more repetitive. Such a finding contradicts the proposition that the valuation of earnings components should increase as these components become more persistent (less transitory). This proposition is even more plausible if repetitive charges are more likely to contain ongoing operating costs, as suggested by both Dechow et al. (1994) and Elliott and Hanna (1996).

Importantly, our analysis complements an examination of how the market valuation of NSIs varies with past frequency by also evaluating how special item persistence varies with frequency.

If the market is valuing NSIs conditional on past frequency, then the observed valuations should correspond with the associated persistence levels. For instance, if persistence increases with prior frequency, then we would expect market valuation to increase with frequency as well. A lack of symmetry between the market valuation of repetitive NSIs and the associated persistence properties would imply the market is either missing or ignoring the persistence implications of special items.

Our analysis also provides some indirect insights about the underlying circumstances that lead to the reporting of NSIs. Such reporting depends on two fundamentals: (1) the occurrence of an unusual or infrequent event that can be classified as a NSI, and (2) management discretion over the reporting of NSIs. Observed NSI frequency for a firm depends on both of these factors. *Ceteris paribus*, firms are more likely to repetitively report special items in operating environments where such events are common. Similarly, management that is more conservative in its special item designations is less likely to report repetitive NSIs than management that is more aggressive in its definition of special items. Consequently, observing that NSI market valuations and persistence effects increase with past frequency is consistent with either the notion that persistent special items are an underlying characteristic of the firm's operating environment or that high frequency is symptomatic of discretionary special item reporting (i.e., shifting ordinary persistent operating items into "transitory" special items).[4]

We attempt to distinguish between these two underlying explanations for special item persistence in two substantive ways. First, we incorporate a number of supplemental control variables that reflect operating environment differences between high- and low-frequency NSI reporting firms. That is, we explicitly control for operating environment factors shown to influence special item reporting across firms. Second, we conduct a supplemental targeted analysis of a specific type of special item—restructuring charges. Nominally, such charges should be purely transitory. Hence, there should be minimal variation in persistence outside managerial misclassification. Consequently, restructuring charges should provide a cleaner setting for evaluating whether manager discretion is an underlying driver of NSI persistence and its related valuation effects.

Persistence of Repetitive Special Items

We examine persistence by starting with the basic persistence relation:

$$INC_{it+j} = \alpha_0 + \alpha_1 ADJ\_INC_{it} + \alpha_2 SI_{it}. \quad (1)$$

INC_{it+j} is quarterly income before extraordinary items for firm i in quarter $t+j$ (Compustat data item #8), where $j = 1$ or 4, scaled by the market value of equity at the beginning of quarter t; RSI_{it} is after-tax quarterly special items (Compustat data item #32) reported by firm i in quarter t;[5] SI_{it} is RSI_{it} scaled by the market value of equity at the beginning of quarter t; and ADJ_INC_{it} is quarterly income before extraordinary items for firm i in quarter t (Compustat data item #8) less RSI_{it} scaled by the market value of equity at the beginning of quarter t. The SI coefficient in Equation (1) measures the average persistence of special items related to next quarter's earnings

[4] Our analysis cannot distinguish whether such discretionary misclassification is intentional or unintentional. For example, a firm may discover, *ex post*, that recurring costs were mistakenly reported as "nonrecurring" items in prior years when the costs should have been included in SG&A. Eastman Kodak is a perfect example of this notion. Kodak has reported a restructuring charge in 11 of the past 15 years. These costs were necessitated by the changing business environment that Kodak operates in and could have been reported as operating expenses. In fact, IBM experienced similar restructuring activity in the 1990s and ultimately announced it would no longer report these costs as restructuring charges because they represented recurring operating costs necessitated by IBM's competitive environment (Ziegler 1996).

[5] Pre-tax values are converted to after-tax values by multiplying the reported pre-tax value by one minus the quarterly effective tax rate measured as total tax expense (Compustat data item #6) divided by pretax income (Compustat data item #23).

(t+1) or earnings four quarters hence (t+4). We also expand Equation (1) to include separate slope coefficients on earnings and special items depending on the frequency of prior reported special items, as follows:

$$INC_{it+j} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \alpha_k ADJ\_INC * FREQ\_k_{it} + \sum_{k=0}^{5} \lambda_k SI * FREQ\_k_{it}. \quad (2)$$

$FREQ\_k_{it}$ equals 1 if firm i reported k special items in the 12 quarters prior to quarter t, where k takes on values of 0 through 5, otherwise 0. When k takes on the value of 5, this indicates that firm i reported at least five special items in the 12 quarters prior to quarter t.[6] In all earnings persistence regressions, the t-statistics we rely on are based on robust-standard errors using two-way clustering by firm and year to control for both time-series and cross-sectional dependence (Gow et al. 2010).[7]

We estimate Equation (2) separately for positive and negative special items, although NSIs are our primary focus. We estimate separate *SI* slope coefficients (λ_0 through λ_5) depending on the number of prior reported special items. If the increased frequency of prior special items signals that these recurring nonrecurring items have a level of persistence that exceeds truly nonrecurring items, then the coefficient on *SI* will be significantly greater when preceded by five or more special items than when no prior special items are reported (i.e., $\lambda_5 > \lambda_0$). We also estimate separate *ADJ_INC* slope coefficients (α_0 through α_5) depending on the number of prior reported special items. If the repetitive reporting of special items negatively influences the persistence of adjusted earnings, then the coefficient on *ADJ_INC* when preceded by five or more special items will be smaller than the coefficient on *ADJ_INC* when preceded by no prior special items (i.e., $\alpha_5 < \alpha_0$).

Market Valuation of Negative Special Items

The valuation aspect of our analysis is, in part, an extension of Elliott and Hanna (1996). Hence, we begin with a replication of their analysis by estimating the following equation:

$$EA\_RET_{it} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \varphi_k \Delta AINC * FREQ\_k_{it} + \sum_{k=0}^{5} \delta_k NSI * FREQ\_k_{it}. \quad (3)$$

EA_RET_{it} is the two-day earnings announcement period [−1,0] market-adjusted stock return (adjusted for value-weighted index) for firm i in quarter t; $RNSI_{it}$ is firm i's tax-adjusted negative special item (Compustat data item #32); NSI_{it} is $RNSI_{it}$ scaled by the market value of equity in quarter t−4;[8] INC_{it-j} is income before extraordinary items (Compustat item #8) for firm i in quarter $t-j$; and $\Delta AINC_{it}$ (unexpected earnings before NSIs) is measured as $(INC_{it} - RNSI_{it} - INC_{it-4}) / MVAL_{it-4}$, where *MVAL* is market value of equity. All other variables are as previously defined. This formulation is a single equation equivalent to the multiple equation method employed by Elliott and Hanna (1996). Based on their analysis, we expect that the earnings response coefficient is higher for firms with no prior NSIs than for firms with one, two or three prior NSIs (i.e., $\varphi_0 > \varphi_1$, $\varphi_0 > \varphi_2$, and $\varphi_0 > \varphi_3$).[9] We also expect that *NSI* coefficients (δ_0 through δ_5) will be smaller than companion earnings response coefficients (φ_0 through φ_5).

---

[6] The measurement of *FREQ* depends on whether the test focuses on positive or negative special items. For example, in our tests of the persistence of NSIs, we measure *FREQ* as the number of prior reported NSIs.

[7] We thank Ian Gow for providing us with the SAS code to perform the two-way clustering procedure.

[8] We define NSI_{it} as any Compustat quarterly data item #32 that exceeds 0.5 percent of total assets at the beginning of the quarter. We perform sensitivity analyses employing the 1.0 percent of total assets cutoff in Elliott and Hanna (1996) as well as all NSIs without limitation. Our results are robust to these alternative specifications.

[9] Elliott and Hanna (1996) find that the earnings response coefficient (ERC) for no write-off observations is significantly greater than the ERC on one, two, or three write-off observations. However, they find no significant difference between the ERC on no write-off observations and the ERC on four and five or more write-off observations.

Elliott and Hanna (1996) examine the two-day announcement period price response to NSIs consistent with their primary focus on how NSIs impact the information content of adjusted earnings. On the other hand, we are directly concerned with the valuation of the special item. Elliott and Hanna (1996, 153) note, "firms reporting large write-offs might be providing more information prior to the earnings announcement." This notion is consistent with prior research detailing how managers pre-announce bad news in an attempt to induce downward revision in analyst forecasts (e.g., Soffer et al. 2000). Material NSIs are often publicly disclosed prior to the earnings announcement date (e.g., Elliott and Shaw 1988; Lopez 2002). In an examination of 1,220 restructuring observations in our sample, we find over 65 percent were announced four or more days prior to the earnings announcement date. Such "bad news" pre-announcements potentially compromise the ability of an earnings-announcement period return design to fully capture the market response to the charge.[10] The use of this short return window potentially biases the coefficient on nonrecurring charges toward zero. In order to ensure that we capture the full valuation impact of the NSI, our primary analysis incorporates a quarterly return similar to Freeman and Tse (1992).[11]

Expanding the return window from a two-day window to a quarterly accumulation period weakens the ability of our analysis to pinpoint the earnings announcement as the direct source of the market price movement, but insulates against the potentially biasing effect of NSI pre-announcements. We employ an equation identical to Equation (3) in all respects except that the dependent variable, $QRET_{it}$, is defined as the quarterly market-adjusted stock return for firm i accumulated over the period from one trading day after the earnings announcement date for the prior quarter ($t-1$) through the earnings announcement date for the current quarter (t). We modify Equation (3) as follows:

$$QRET_{it} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \varphi_k \Delta AINC * FREQ\_k_{it} + \sum_{k=0}^{5} \delta_k NSI * FREQ\_k_{it}. \quad (3a)$$

Since value-relevant earnings information commonly impacts returns prior to the earnings announcement (Kothari and Sloan 1992), the estimated coefficients for Equation (3a) are likely to be higher than those for Equation (3). Consistent with Elliott and Hanna (1996), we anticipate that, as the frequency of prior reported NSIs increases, the decline in the $\Delta AINC$ coefficients should remain intact (i.e., $\varphi_0 > \varphi_1$, $\varphi_0 > \varphi_2$, and $\varphi_0 > \varphi_3$) and the coefficient on $\Delta AINC$ will be significantly greater than the coefficient on NSI when $k = 0$ (no prior NSIs). However, we have no *a priori* expectation of the relation between unexpected earnings ($\Delta AINC$) and special items (NSI) as the sequence of reported NSIs increases. Finally, if the market perceives an increase in special items reporting frequency as an increase in the persistence effects of these items, then the coefficient on NSI will be increasing in k (i.e., $\delta_5 > \delta_0$).

Market Valuation of Restructuring Charges

Our final analysis focuses on a subset of special items identified as restructuring charges. We evaluate how the market values adjusted income and restructuring charges using the following variations of Equations (3) and (3a):

---

[10] This caveat does not apply to an evaluation of the response to an earnings "surprise" contained in the earnings announcement. If the coefficient on income (earnings surprise) is of primary importance, as is the case in Elliott and Hanna (1996), then the announcement period analysis is highly appropriate.

[11] Freeman and Tse (1992) select a quarterly return window in their analysis because they did not have the exact date of analysts' forecasts. They indicate that use of a shorter window that excludes a portion of the market reaction to the analysts' forecasts will result in a downward biased ERC (Freeman and Tse 1992, 189).

$$EA\_RET_{it} = \sum_{k=0}^{4} \beta_k FREQ\_k_{it} + \sum_{k=0}^{4} \theta_k \Delta AINC\_RC{*}FREQ\_k_{it} + \sum_{k=0}^{4} \gamma_k RC{*}FREQ\_k_{it} \quad (4)$$

and

$$QRET_{it} = \sum_{k=0}^{4} \beta_k FREQ\_k_{it} + \sum_{k=0}^{4} \theta_k \Delta AINC\_RC{*}FREQ\_k_{it} + \sum_{k=0}^{4} \gamma_k RC{*}FREQ\_k_{it}. \quad (4a)$$

RRC_{it} is firm i's tax-adjusted restructuring charge; $\Delta AINC\_RC_{it}$ is unexpected earnings before restructuring charges for firm i in quarter t, measured as $(INC_{it} - RRC_{it} - INC_{it-4}) / MVAL_{it-4}$, where $MVAL$ is market value of equity; RC_{it} equals RRC_{it} scaled by the market value of equity in quarter t–4; and $FREQ\_k_{it}$ equals 1 if firm i reported k restructuring charges in the 12 quarters prior to quarter t, where k takes on values of 0 through 4. When k takes on the value of 4, this indicates that firm i reported at least four restructuring charges in the 12 quarters prior to quarter t. All other variables are as previously defined.

The behavior of compensation committees with respect to repetitive restructuring charges documented in Dechow et al. (1994) and Adut et al. (2003) implies that $\gamma_0 < \gamma_1 < \gamma_2 < \gamma_3 < \gamma_4$ in the estimation of Equation (4a). That is, we expect that the response coefficient on restructuring charges will increase with observed frequency of prior reported restructurings. For θ_0 through θ_4, we have no reason to anticipate any substantive departure from the findings of Elliott and Hanna (1996). Hence, we expect these coefficient values to be positive but declining in magnitude as the frequency of prior reported restructuring charges increases.

IV. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

Our analysis employs three separate samples: (1) to estimate the persistence of special items, (2) to estimate the market valuation of repetitive special items, and (3) to estimate the market valuation of repetitive restructuring charges, a more homogenous subset of NSIs. For our first two analyses, we employ the Compustat Industrial Quarterly database to identify the special item sample. The sample of special item observations consists of firms reporting at least one special item (Compustat data item #17) between 1992 and 2003. From this set, we identify all firm-quarters with available Compustat and CRSP data necessary to estimate the regression equations. The earnings persistence tests in Equation (2) require complete data for 17 quarters surrounding quarter t (t–12 through t+4), the quarter in which the special item is recognized. We then eliminate as an outlier any observation that exceeds the cutoffs proposed in Belsley et al. (1980) for at least three of four outlier metrics (hat matrix, studentized residual, COVRATIO statistic, and dfFITS statistic). Our data requirements and sample selection criteria result in a test sample of 112,721 firm-quarter observations for the estimation of Equation (2). To perform our market valuation tests, Equations (3) and (3a), we identify all firm-quarter observations from 1992 to 2003 with available Compustat and CRSP data necessary to estimate the corresponding equations and eliminate outliers consistent with our earnings persistence tests noted above. This process results in test samples of 151,532 and 151,498 firm-quarter observations to estimate Equations (3) and (3a), respectively.[12]

The restructuring observations in our sample are all associated with corporate downsizing. While not perfectly homogeneous, we limit our restructuring observations to downsizings because we believe the similarity of an event is essential when seeking to understand how the market

[12] The difference in sample sizes for Equations (3) and (3a) are due to our method of addressing outliers. We also ran our tests using a 1 percent cutoff for all regression variables (resulting in the same sample size for these tests) and our primary results are robust to this alternative method of addressing outliers.

responds to a series of sequential events. Our sample of restructuring observations consists of firms reporting at least one restructuring charge between 1992 and 2003. For years prior to 1995, restructuring observations are identified by searching the National Automated Accounting Research System (NAARS) using the search strings "restruct!," "unusual," and "special." For years between 1994 and 1998, the Lexis-Nexis Academic Universe Business News database is searched using the same search strings. We then review the annual financial statements and footnotes of each firm identified in this search to assure each firm actually reported a restructuring charge related to downsizing. This search and review yields an initial sample of 642 "restructuring" firms. To identify quarterly restructuring observations for this sample of 642 firms, we review the 10-Q reports of these firms for every quarter between 1992 and 2003. From this search, an initial sample of 2,278 quarterly restructuring observations is identified.[13]

We estimate cross-sectional regression models using all quarterly restructuring charges, returns, and earnings data available for the period from 1992 through 2003 (a maximum of 48 time-series observations for each firm). The firm must report at least one quarterly restructuring charge from 1992 to 2003 and have Compustat and CRSP data necessary to estimate all equations. These limitations result in a test sample of 1,220 quarterly restructuring observations taken by 382 firms for purposes of estimating Equations (4) and (4a) for a total of 15,533 firm-quarters.

Panel A of Table 1 reports descriptive statistics for the full special item sample. Mean (median) total assets and market value of equity are $2.9 billion ($215 million) and $1.9 billion ($161 million), respectively. The mean (median) negative special item is 3.2 (0.7) percent of market value while the mean (median) positive special item is 2.2 (0.4) percent of market value. Average quarterly (earnings announcement) market-adjusted return is 0.37 percent (0.43 percent). The mean quarterly effective tax rate (ETR) is 25 percent and the mean operating margin is 1.9 percent. Finally, our sample firms have an average of 1.8 reportable operating segments.[14]

Panel B of Table 1 reports mean descriptive statistics for those firms that report a NSI conditioned on the frequency of prior reported NSIs. More than 16 percent of our NSI observations report no NSIs the prior 12 quarters (No Prior NSI), while more than a third report four or more prior NSIs. The No Prior NSI observations generally experience lower returns (both the quarterly and short windows, *QRET* and *EA_RET*, respectively), are smaller, and report lower income than observations with multiple prior NSI incidences.[15] In fact, income (ΔINC and $\Delta AINC$) and market return (both *QRET* and *EA_RET*) means for all groups in Panel B are significantly lower (not tabulated) than the overall sample means reported in Panel A, consistent with prior research showing that firms are more likely to report/experience special items during periods of poor performance (e.g., Elliott and Shaw 1988; DéAngelo et al. 1994; Carter 2000). However, Panel B suggests a performance improvement as prior NSI frequency increases. That is, as prior NSI frequency increases, ΔINC, *QRET*, and *EA_RET* tend to become less negative, while the two adjusted income measures ($\Delta AINC$ and *OPMARGIN*) become positive. Hence, while NSIs are associated with poor performance, we find that performance improves toward a more normal level as the frequency of reporting NSIs increases. The improved (or less negative) performance level

[13] Our search for restructuring firms is limited to the years 1992 through 1998. We then search the 10-Qs for this sample for every quarter from Q1 1992 through Q4 2003 to identify quarterly restructuring observations for the 642 firms identified in our initial search. The primary reason we stopped our search at 1998 is the significant amount of hand-collection of data required.

[14] In an untabulated analysis, we estimate descriptive statistics for our restructuring sample. In general, these statistics are quantitatively and qualitatively similar to those reported here. The one noteworthy difference is restructuring firms are larger. Mean (median) total assets for the restructuring sample is $5.1 billion ($482 million).

[15] Differences between the More than four NSI means and the No Prior NSI means in Panel B are significant at the 0.05 level or higher (not tabulated) for all variables.

TABLE 1
Descriptive Statistics

Panel A: Full Sample

	n[a]	Mean	Standard Deviation	Lower Quartile	Median	Upper Quartile
QRET (percentage)	151,688	0.367	0.284	−14.437	−0.186	14.024
EA_RET (percentage)	151,688	0.430	0.067	−6.776	0.090	3.188
ASSETS ($million)	151,688	2,889	21,230	54	215	895
MVAL ($million)	151,688	1,910	11,485	42	161	705
NSI (percentage)[b]	23,055	−3.186	0.074	−2.616	−0.692	−0.192
PSI (percentage)[b]	8,328	2.214	0.059	0.115	0.414	1.471
ETR (percentage)	151,688	24.991	0.202	1.384	33.529	38.484
Δ*INC* (percentage)	151,688	0.273	0.056	−0.658	0.188	0.990
Δ*AINC* (percentage)	151,688	0.636	0.049	−0.517	0.215	1.031
OPMARGIN (percentage)	151,688	1.884	0.758	−13.433	12.441	23.409
SEGMENTS	151,688	1.778	1.389	1	1	3

Panel B: Mean Values Where *SI* < 0 in Quarter *t* Conditioned on Number of Prior Negative Special Items

	No Prior NSI[c]	1 Prior NSI	2 Prior NSI	3 Prior NSI	4 Prior NSI	More Than 4 NSI
# of Quarterly Observations	3,742	4,438	3,971	3,137	2,371	5,396
QRET (percentage)	−3.618	−2.603	−2.668	−2.153	−1.528	−2.101
EA_RET (percentage)	−0.054	0.042	−0.254	0.013	0.148	0.263
ASSETS ($million)	2,617	3,110	2,902	3,530	5,195	10,940
MVAL ($million)	1,632	2,025	2,070	2,579	3,280	7,904
NSI (percentage)	−3.563	−3.493	−3.451	−3.250	−3.128	−2.464
ETR (percentage)	23.794	21.326	20.301	19.676	19.604	19.068
Δ*INC* (percentage)	−3.623	−3.071	−2.533	−2.073	−1.415	−0.259
Δ*AINC* (percentage)	−0.515	−0.071	0.370	0.637	1.340	1.808
OPMARGIN (percentage)	−4.814	−7.522	−6.553	−5.382	−4.065	2.280
SEGMENTS	1.825	1.924	1.943	2.014	2.064	2.180

(continued on next page)

Panel C: Mean Values Where $SI > 0$ in Quarter t Conditioned on Number of Prior Positive Special Items

	No Prior PSI[c]	1 Prior PSI	2 Prior PSI	3 Prior PSI	4 Prior PSI	More Than 4 PSI
# of Quarterly Observations	3,213	2,142	1,338	727	408	500
QRET (percentage)	2.007	2.464	0.627	1.221	1.432	−1.485
EA_RET (percentage)	0.885	0.961	0.373	0.658	0.446	0.357
ASSETS ($million)	3,135	4,521	4,858	5,609	6,585	6,823
MVAL ($million)	2,149	3,786	4,382	4,451	4,899	5,890
PSI (percentage)	2.315	2.224	1.949	2.226	2.089	2.317
ETR (percentage)	24.806	24.585	25.124	25.229	26.007	25.363
Δ*INC* (percentage)	3.088	2.890	2.495	2.479	2.196	1.069
Δ*AINC* (percentage)	1.147	1.138	0.892	0.930	0.552	−0.660
OPMARGIN (percentage)	−0.021	3.062	4.953	8.650	8.360	7.320
SEGMENTS	1.881	1.878	1.942	1.971	1.946	1.774

[a] n = number of quarterly observations.

[b] Negative (positive) special items only include observations where the quarterly special item is less than (greater than) 0.

[c] Prior recognized negative and positive special items are measured over the prior 12 quarters.

Variable Definitions:

$QRET_{it}$ = quarterly market-adjusted stock return for firm i (adjusted for value-weighted index) accumulated over the period from one trading day after the earnings announcement date for the prior quarter $(t-1)$ through the earnings announcement date for the current quarter (t);

EA_RET_{it} = two-day earnings announcement period $[-1,0]$ market-adjusted stock return (adjusted for value-weighted index) for firm i in quarter t;

$ASSETS_{it}$ = total assets for firm i in quarter t in millions;

$MVAL_{it}$ = market value of equity for firm i in quarter t in millions;

SI_{it} = the after-tax amount of firm i's special item for quarter t (Compustat data item #32) divided by the market value of equity in quarter $t-4$ times 1 minus the quarterly effective tax rate (ETR_{it});

$NSI_{it} = SI_{it}$ if $SI_{it} < 0$, otherwise 0;

$PSI_{it} = SI_{it}$ if $SI_{it} > 0$, otherwise 0;

ETR_{it} = effective tax rate for firm i in quarter t (Compustat data item 6) divided by pre-tax income (Compustat data item 23);

ΔINC_{it} = unexpected earnings for firm i in quarter t scaled by the market value of equity. Measured as income before extraordinary items (Compustat data item #8) in quarter t less income before extraordinary items in quarter $t-4$ divided by the market value of equity in quarter $t-4$;

$\Delta AINC_{it} = INC_{it}$ less SI_{it};

$OPMARGIN_{it}$ = operating income before depreciation (Compustat data item #21) scaled by sales (Compustat data item #2) for firm i in quarter t; and

$SEGMENTS$ = number of reportable segments for firm i in quarter t.

for high-frequency NSI firms is broadly consistent with the notion that these charges reflect ongoing or more "normal" activities of the firms in this category.

Firm size also increases with NSI frequency as mean assets and market value of the More than 4 NSI group are approximately four times those of the No Prior NSI group. One reason for such a relation is that larger firms tend to report more segments than smaller firms. In multi-segment firms, special item disclosures are often determined on the basis of whether a charge is material at the segment level. *Ceteris paribus,* the more segments a firm reports, the more likely the firm will report a special item. The increase in average number of segments from 1.825 for the No Prior NSI group to 2.180 for the More than 4 NSI group is consistent with the number of segments contributing to the frequency of NSI reporting.

As prior reporting frequency increases, the average effective tax rate (ETR) decreases monotonically from 23.8 percent down to 19.1 percent. Our main results are based on tax-adjusted NSIs, and may be sensitive to the tax rate adjustment. In sensitivity analyses, we find our results to be robust to the use of pre-tax NSI values as well as to several alternative tax rate adjustment methods. Finally, the magnitude of NSIs (measured on either a pre-tax or tax-adjusted basis) declines monotonically as prior reporting frequency increases with the sharpest decline at the More than 4 prior NSI category. Such a decline in magnitude suggests that firms repeatedly reporting NSIs may have a lower materiality threshold for NSIs than firms that report less frequently. In addition, it is also consistent with the notion that some firms that frequently report NSIs are engaging in a planned multi-period event where only part of the total cost of an event is recognized in any one period.

Panel C of Table 1 reports mean descriptive statistics for those firms that report a positive special item (*PSI*) conditioned on the frequency of prior reported PSIs. There are marked differences between the positive and negative special item subsamples. There are noticeably fewer recurring PSIs as approximately 39 percent report no PSIs in the prior 12 quarters while only 11 percent report four or more prior PSIs. The frequency-related decrease noted in ETRs for NSIs is not found as the ETR for PSIs remains relatively stable at 25 percent, regardless of prior reporting frequency. Unlike the NSI sample, we find that adjusted income decreases with the number of prior reported PSIs. Similar to the NSI sample, firm size increases with the frequency of prior reported PSIs, but no such relationship is apparent for the number of reportable segments.

In a supplemental descriptive analysis reported in Appendix A, we examine NSI properties for a randomly selected subsample of 100 high-frequency NSI observations (three or more NSIs over the prior 12 quarters). This analysis indicates that restructuring charges are the most common, but not dominant, type of NSI reported in this subsample at 34 percent of the total.[16] The remaining reported items show the considerable degree of heterogeneity in the cross-section of high-frequency special items that makes measurement of repeated *SI* effects somewhat problematic. This analysis also shows that a considerable degree of heterogeneity exists within the time-series of charges for these firms. Merger-related costs is the only category where over time frequency rates approach 50 percent and these rates are under 30 percent for most other special item types.

Appendix B reports prior reporting frequencies of NSIs by industry classification. Industries with the greatest frequency of NSI recognition are business services, electronic equipment, computers, banking, and pharmaceuticals, respectively. These five industries represent over 40 percent of all quarterly NSI observations in our sample. Because there is evidence of substantial industry clustering, we control for possible industry effects in our empirical testing.

[16] The special item types are restructuring charges, merger-related costs, in-process R&D, asset write-offs, litigation costs, goodwill write-off, asset disposition, and other.

V. RESULTS

Persistence Properties of Negative Special Items

To begin, we explore the persistence properties of NSIs. Specifically, we investigate whether the frequency of prior special item recognition relates to the frequency of subsequent recognition. We perform this analysis on negative and positive special item subsamples. For the NSI subsample, 16,430 observations report at least one NSI during the test period and have complete data for the 25 quarters from $t-12$ to $t+12$. For the PSI subsample, 5,294 observations report at least one PSI during our test period and have complete data for the 25 quarters from $t-12$ to $t+12$.

Panel A of Table 2 reports results for the NSI analysis. On average, sample observations report 2.9 NSIs over the 12 quarters subsequent to quarter t. For NSIs, the results suggest the frequency of prior reporting is significantly related to the frequency of subsequent reporting. The mean number of subsequent NSIs for observations that are preceded by no NSIs over the prior 12 quarters is 1.9. On the other hand, for observations preceded by five or more prior reported NSIs, the mean number of subsequent charges is 4.5. This 2.7 difference in subsequent frequency is significant at the 0.01 level. We also investigate the magnitude of subsequent NSIs conditioned on the frequency of prior reported charges. We find that, for observations with no prior reported NSIs, the aggregate charges over the subsequent 12 quarters represent an average of 7.6 percent of the market value of equity in quarter t. However, if the current special item is preceded by five or more prior NSIs, the aggregate charges over the subsequent 12 quarters represent an average of 14.9 percent of the market value of equity in quarter t. This 7.3 percent difference is significant at the 0.01 level. Finally, for observations with no prior reported NSIs, the aggregate charges over the subsequent 12 quarters represent 198 percent of the charge taken in quarter t. However, if the current special item is preceded by five or more prior NSIs, the aggregate charges over the subsequent 12 quarters represent 357 percent of the charge reported in quarter t. This 159 percentage point difference is also significant at the 0.01 level. The evidence suggests that both the frequency and magnitude of subsequently reported NSIs are significantly associated with past reporting frequency.

Panel B of Table 2 reports results for the PSI analysis. PSI firms report 1.5 PSIs over the 12 quarters subsequent to quarter t, approximately half of that reported for NSIs. The mean number of subsequent PSIs is 1.2 (2.3) for observations that are preceded by no (five or more) PSIs over the prior 12 quarters. This difference of 1.1 is significant at the 0.01 level; however, it is considerably smaller than the NSI difference of 2.7. The aggregate value of PSIs over the subsequent 12 quarters represents 3.6 (5.9) percent of the market value of equity in quarter t for observations that are preceded by no (five or more) PSIs over the prior 12 quarters. This 2.3 percent difference is marginally significant ($p < 0.10$) and less than a third of the NSI difference of 7.3. Finally, the aggregate value of PSIs recognized over the subsequent 12 quarters represents 136 (252) percent of the PSI reported in quarter t for observations that are preceded by no (five or more) PSIs over the prior 12 quarters. This 116 percentage point difference is significant at the 0.01 level, but again substantially less than the NSI difference of 159 percent. Regardless of sign, the evidence suggests that both the frequency and magnitude of subsequently reported special items are significantly associated with past reporting frequency, but the association is stronger for NSIs.

Table 3 reports estimates of Equation (2), which measures the persistence of current earnings and special items related to next quarter earnings or subsequent seasonal earnings (fourth subsequent quarter). Panel A reports NSI estimates, while Panel B reports PSI estimates. Table 3 also provides estimates for an augmented version of Equation (2), which introduces controls for empirically identified determinants of repetitive special item frequencies. The descriptive evidence reported in Table 1, Panel B as well as prior literature (e.g., Francis et al. 1996; Bens and Johnston 2009) suggest that several factors are significantly associated with the recognition of a negative

TABLE 2

Frequency of Subsequent Special Item Recognition, Magnitude of Subsequent Special Items, and the Ratio of the Current Special Item to Subsequent Special Items Conditioned on the Frequency of Prior Reported Special Items

Panel A: Negative Special Items

	All Observations (n = 16,430)	No Prior Reported NSI (n = 2,891)	1 Prior Reported NSI (n = 3,191)	2 Prior Reported NSI (n = 2,855)	3 Prior Reported NSI (n = 2,194)	4 Prior Reported NSI (n = 1,675)	5 or More Prior Reported NSI (n = 3,624)	Difference 5 or More Less No Prior NSI
SUB_FREQ	2.874	1.859	2.373	2.621	3.022	3.635	4.536	2.678***
SUB_NSI%	−12.090	−7.590	−11.254	−12.193	−14.572	−15.538	−14.898	−7.309***
SUB_CURR_N%	267.8	198.1	240.1	265.4	291.6	301.6	357.2	159.0***

Panel B: Positive Special Items

	All Observations (n = 5,294)	No Prior Reported PSI (n = 2,294)	1 Prior Reported PSI (n = 1,601)	2 Prior Reported PSI (n = 948)	3 Prior Reported PSI (n = 548)	4 Prior Reported PSI (n = 335)	5 or More Prior Reported PSI (n = 354)	Difference 5 or More Less No Prior NSI
SUB_FREQ	1.478	1.211	1.450	1.612	1.707	2.060	2.337	1.125***
SUB_PSI%	4.524	3.647	4.274	5.498	6.232	5.610	5.917	2.270*
SUB_CURR_P%	169.1	136.1	171.0	189.6	200.5	221.7	251.7	115.6***

*, **, *** Two-tailed $p < 0.10$, two-tailed $p < 0.05$, and two-tailed $p < 0.01$, respectively.

The analysis for this table requires complete data for all items for the 25 quarters surrounding the quarter *t*, the quarter in which the positive or negative special item is recognized.

Variable Definitions:

SUB_FREQ = number of quarterly negative (positive) special items recognized by firm *i* over the 12 quarters subsequent to quarter *t* if a negative (positive) special item was recognized in quarter *t*;

SUB_NSI% = sum of negative special items recognized by firm *i* over the 12 quarters subsequent to quarter *t* if a negative special item was recognized in quarter *t* scaled by the market value of equity in quarter *t* stated in percentage terms;

SUB_PSI% = sum of positive special items recognized by firm *i* over the 12 quarters subsequent to quarter *t* if a positive special item was recognized in quarter *t* scaled by the market value of equity in quarter *t* stated in percentage terms;

SUB_CURR_N% = *SUB_NSI* divided by *NSI* where *NSI* equals the negative special item recognized in quarter *t* divided by the market value of equity in quarter *t* stated in percentage terms; and

SUB_CURR_P% = *SUB_PSI* divided by *PSI* where *PSI* equals the positive special item recognized in quarter *t* divided by the market value of equity in quarter *t* stated in percentage terms.

TABLE 3

Results of the Estimation of Earnings Persistence Equation (2)

Panel A: Negative Special Items

$$INC_{it+j} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \alpha_k ADJ\_INC * FREQ\_k_{it} + \sum_{k=0}^{5} \lambda_k SI * FREQ\_k_{it} \quad \text{Equation (2)}$$

Variable	Predicted Sign	INC_{it+1} No Controls (n = 112,721)	INC_{it+1} With Controls[b] (n = 74,982)	INC_{it+4} No Controls (n = 112,721)	INC_{it+4} With Controls[b] (n = 74,982)
ADJ_INC_0	+	0.608	0.062	0.537	0.001
		[0.0001][a]	[0.0001]	[0.0001]	[0.9544]
ADJ_INC_1	+	0.595	0.021	0.502	−0.036
		[0.0001]	[0.0388]	[0.0001]	[0.0012]
ADJ_INC_2	+	0.572	0.015	0.449	−0.099
		[0.0001]	[0.2188]	[0.0001]	[0.0001]
ADJ_INC_3	+	0.577	−0.024	0.486	−0.059
		[0.0001]	[0.1227]	[0.0001]	[0.0003]
ADJ_INC_4	+	0.587	0.025	0.414	−0.077
		[0.0001]	[0.1948]	[0.0001]	[0.0001]
ADJ_INC_5	+	0.580	0.032	0.422	−0.055
		[0.0001]	[0.0440]	[0.0001]	[0.0010]
NSI_0	?	−0.065	0.120	−0.022	0.244
		[0.0137]	[0.0001]	[0.0373]	[0.0001]
NSI_1	?	−0.052	0.089	−0.024	0.245
		[0.0138]	[0.0001]	[0.0236]	[0.0001]
NSI_2	?	−0.012	0.110	−0.013	0.259
		[0.4275]	[0.0001]	[0.2891]	[0.0001]
NSI_3	?	−0.012	0.089	0.009	0.297
		[0.6347]	[0.0004]	[0.5705]	[0.0001]
NSI_4	+	0.041	0.178	0.024	0.291
		[0.0321]	[0.0001]	[0.0087]	[0.00001]
NSI_5	+	0.041	0.219	0.029	0.357
		[0.0115]	[0.0001]	[0.0164]	[0.0001]

(continued on next page)

Panel B: Positive Special Items

$$INC_{it+j} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \alpha_k ADJ\_INC * FREQ\_k_{it} + \sum_{k=0}^{5} \lambda_k SI * FREQ\_k_{it} \quad \textbf{Equation (2)}$$

Variable	Predicted Sign	INC_{it+1} No Controls (n = 112,721)	INC_{it+1} With Controls[b] (n = 63,769)	INC_{it+4} No Controls (n = 112,721)	INC_{it+4} With Controls[b] (n = 63,769)
ADJ_INC_0	+	0.579	−0.001	0.502	−0.001
		[0.0001][a]	[0.9915]	[0.0001]	[0.9897]
ADJ_INC_1	+	0.585	0.062	0.459	−0.007
		[0.0001]	[0.3916]	[0.0001]	[0.8981]
ADJ_INC_2	+	0.534	−0.043	0.456	−0.035
		[0.0001]	[0.4426]	[0.0001]	[0.5554]
ADJ_INC_3	+	0.551	−0.003	0.414	−0.146
		[0.0001]	[0.9708]	[0.0001]	[0.0138]
ADJ_INC_4	+	0.465	−0.081	0.343	−0.194
		[0.0001]	[0.1420]	[0.0001]	[0.0314]
ADJ_INC_5	+	0.442	0.003	0.481	0.149
		[0.0001]	[0.9715]	[0.0001]	[0.2516]
PSI_0	?	0.015	−0.080	−0.148	−0.144
		[0.6236]	[0.5924]	[0.0116]	[0.2496]
PSI_1	?	0.135	−0.075	−0.016	−0.059
		[0.0718]	[0.5937]	[0.7491]	[0.5737]
PSI_2	?	−0.023	−0.146	−0.123	−0.073
		[0.5128]	[0.2507]	[0.0391]	[0.6450]
PSI_3	?	0.147	0.109	−0.040	0.016
		[0.0445]	[0.3081]	[0.3676]	[0.8629]
PSI_4	+	0.008	−0.019	0.182	0.072
		[0.8770]	[0.8416]	[0.0045]	[0.5906]
PSI_5	+	0.052	0.179	−0.049	−0.068
		[0.4604]	[0.3225]	[0.6536]	[0.6451]

[a] Two-tailed p. t-statistics are based on robust-standard errors using two-way clustering by firm and year to control for both time-series and cross-sectional dependence (Gow et al. 2010).

(continued on next page)

[b] For the tests reported in columns 4 and 6 we include a vector of controls variables that are included as main effects and interacted with each of the income and special item variables. See Appendix C for definitions of the control variables.

Variable Definitions:

INC_{it+j} = quarterly income before extraordinary items for firm i in quarter $t+j$ (Compustat data item #8), where $j = 1$ or 4, divided by the market value of equity in quarter t;

RSI_{it} = quarterly special items (Compustat data item #32) reported by firm i in quarter t multiplied by 1 minus the quarterly effective tax rate, where the effective tax rate is measured as total tax expense (Compustat data item #6) divided by pretax income (Compustat data item #23);

SI_{it} = RSI_{it} divided by the market value of equity at the beginning of quarter t;

NSI_{it} = SI_{it} if $SI_{it} < 0$, otherwise 0;

PSI_{it} = SI_{it} if $SI_{it} > 0$, otherwise 0;

$FREQ\_k_{it}$ = 1 if firm i reported k special items in the 12 quarters prior to quarter t, where k takes on values of 0 through 5, otherwise 0. When k takes on the value of 5 this indicates that firm i reported at least five special items in the 12 quarters prior to quarter t;

ADJ_INC_{it} = quarterly income before extraordinary items for firm i in quarter t (Compustat data item #8) less RSI_{it}, divided by the market value of equity at the beginning of quarter t;

$ADJ\_INC\_k_{it}$ = $ADJ\_INC_{it} * FREQ\_k_{it}$, where $FREQ\_k$ is based on prior NSI or PSI reporting;

NSI_k_{it} = $NSI_{it} * FREQ\_k_{it}$, where $FREQ\_k$ is based on prior NSI reporting;

PSI_k_{it} = $PSI_{it} * FREQ\_k_{it}$, where $FREQ\_k$ is based on prior PSI reporting; and

$CONTROLS$ = see Appendix C for definitions of the control variables.

special item/restructuring charge. In a supplemental analysis (not tabulated), we identify factors associated with negative and positive special item reporting frequencies.[17] We find that firm size, accounts receivable turnover, PP&E turnover, change in ROE, change in ROA, industry growth in sales, change in industry ROA, prior year annual return, prior five-year annual return, average five-year change in the book-to-market ratio, average five-year change in the industry book-to-market ratio, and an indicator variable for the presence of a loss in any of the three prior years are significant predictors of NSI frequency. We also find that firm size, profit margin, change in ROA, industry growth in sales, change in industry ROA, prior year annual return, prior five-year annual return, average five-year change in the book-to-market ratio, average five-year change in the industry book-to-market ratio, and an indicator variable for the presence of a loss in any of the three prior years are significant predictors of PSI frequency. Our empirical specification measures these control variables as decile ranks and includes them as both main effects and interacted variables with *ADJ_INC* and either *NSI* (Panel A) or *PSI* (Panel B).[18]

The third and fourth (fifth and sixth) columns of Table 3, Panel A report estimations of Equation (2) with quarter t+1 (t+4) earnings as the dependent measure, without and with control variables, respectively. The results without controls are broadly similar for both future earnings designations (t+1 and t+4). The coefficient on *ADJ_INC* generally decreases as the frequency of prior NSIs increases. For t+1 earnings, the *ADJ_INC_0* coefficient (no prior NSIs) is 0.608 which is significantly greater ($p < 0.01$) than the 0.580 coefficient on *ADJ_INC_5*. For t+4 earnings, the *ADJ_INC_0* coefficient of 0.537 is also significantly greater than the *ADJ_INC_5* coefficient of 0.422 ($p < 0.01$). When frequency-related control variables are introduced, the income coefficient magnitudes drop as they now reflect marginal rather than total earnings effects. For the t+1 and t+4 earnings regressions, the largest income coefficients continue to be *ADJ_INC_0* (coefficient estimates of 0.062 and 0.001, respectively). In the case of t+4 earnings, the *ADJ_INC_5* coefficient of −0.055 is significantly smaller ($p < 0.01$) than the *ADJ_INC_0* coefficient. In the case of t+1 earnings the difference in these coefficients lacks significance at conventional levels ($p = 0.142$). However, the intervening coefficient estimates (i.e., *ADJ_INC_1* through *ADJ_INC_4*) are all significantly smaller ($p < 0.01$) than the *ADJ_INC_0* coefficient. Collectively, these results suggest that earnings become less persistent as the prior reporting frequency of NSIs increases.

With respect to NSIs, the *NSI_0* coefficients are significantly smaller ($p < 0.05$ or better) than the *NSI_5* coefficients in all four estimations. In the two no-controls estimations, the *NSI_0* coefficients are significantly negative ($p < 0.05$), suggesting that current NSIs not preceded by prior NSIs lead to higher earnings in the subsequent quarter. In contrast, these same coefficients are positive and significant ($p < 0.01$) in the two estimations that include control variables. Thus, after controlling for the determinants of NSIs, the marginal effect of a current NSI is a decrease in future earnings. Regardless of whether it is a main effect or marginal effect, the evidence remains consistent with the notion that the persistence of NSIs with respect to future earnings increases with prior reporting frequency.[19]

Table 3, Panel B reports estimations of Equation (2) for PSIs. The results with respect to the earnings/frequency interaction terms when estimating without controls are generally consistent

[17] The supplemental analysis is available from the second author upon request. The variables included in this analysis are defined in Appendix C.

[18] We also use continuous versions of these controls in additional testing and the results are quantitatively and qualitatively similar to the tabulated results.

[19] We estimate seasonal earnings difference variations of Equations (1) and (2). The results of these additional tests (not tabulated) are quantitatively and qualitatively similar to those tabulated in Table 3. In addition, we used two-way clustering by industry and year to control for the industry clustering we note in Appendix B. Our primary conclusions are unaffected.

with the results reported in Panel A. That is, the persistence of earnings decreases with increased reporting frequency of PSIs. In general, once the control variables are included, earnings/frequency interaction terms become insignificant, suggesting that PSI reporting frequency has little or no marginal effect on earnings persistence. However, unlike the results in Panel A for NSIs, the PSI results provide no evidence of a relation between the prior reporting frequency of PSIs and their persistence effects on future earnings. Consistent with this conclusion, we find no clear pattern in the *PSI_k* interaction terms based on frequency. In fact, regardless of the model we estimate (i.e., quarterly or seasonal persistence with or without control variables), the *PSI_k* variables are generally insignificant at conventional levels.

Our results with respect to frequency of prior NSIs and earnings persistence triangulate nicely with the market results of Elliott and Hanna (1996). The persistence of earnings before NSIs (our result) and the market valuation multiple on earnings before NSIs (their result) both decrease with the prior reporting frequency of NSIs. In contrast, the Elliott and Hanna (1996) results suggest the valuation multiple on NSIs decreases with prior frequency, whereas our analysis suggests that future earnings implications of NSIs become more persistent. The increased persistence of NSIs equates to lower future earnings, since the special items are entered as negative amounts. In addition to the persistence effects, our Table 2 results suggest that increased prior reporting frequency is associated with a higher probability of reporting future NSIs (both in frequency and magnitude). In light of the increased persistence characteristics and the future income-decreasing implications of repetitive NSIs documented in Tables 2 and 3, the Elliott and Hanna (1996) results with respect to the market valuation of repetitive charges become difficult to reconcile. Is the market fooled by repetitive NSIs? Or is the market somehow missing the future earnings implications of repetitive NSIs? Alternatively, is there something missing in the market valuation models of prior research? We investigate the latter possibility by considering several alternative specifications of the Elliott and Hanna (1996) model.

Market Valuation of Repetitive Negative Special Items

Table 4, Panel A reports regression results from the estimation of Equations (3) and (3a) for NSIs.[20] Panel B reports coefficient difference tests from the estimation of Equations (3) and (3a). We also estimate expanded versions of Equations (3) and (3a), which include controls for other factors associated with the reporting frequency of NSIs (see Appendix C for control variable definitions). In the Equation (3) short-window earnings announcement period return analysis (column 3, No Controls), our findings are generally consistent with those reported by Elliott and Hanna (1996). The series of $\Delta AINC\_k$ coefficients declines from a high of 0.233 for firm-quarters with no prior NSIs (in the 12 quarters prior to the current quarter) to 0.099 for firm-quarters with five or more prior NSIs. A similar pattern also holds after inclusion of the control variables and associated interactions (column 4). The coefficient on $\Delta AINC\_0$ (no prior NSIs) is significantly greater ($p < 0.01$) than each of the other $\Delta AINC\_k$ coefficients in either *EA_RET* model (test results not tabulated).

With respect to the short-window valuation of NSIs without controls (Equation (3)), the coefficients on *NSI_0* through *NSI_4* each significantly exceed zero ($p < 0.06$ or better), but the coefficient on *NSI_5* is insignificant. The *NSI_0* coefficient of 0.039 exceeds the *NSI_5* coefficient of 0.001, consistent with a decline in the weight the market attaches to NSIs as the sequence of these items increases, as reported by Elliott and Hanna (1996). However, the difference between

[20] Given the comparison to Elliott and Hanna (1996) and our Table 3 results that show little or no relation between the prior reporting of PSIs and the earnings persistence effects of PSIs, we focus the remainder of our analyses solely on NSIs.

TABLE 4

Cross-Sectional Regressions of Earnings Announcement and Quarterly Market-Adjusted Returns on Unexpected Earnings before Negative Special Items and Negative Special Items

Panel A: Estimation of Equations (3) and (3a)

$$R_{it} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \phi_k \Delta AINC * FREQ\_k_{it} + \sum_{k=0}^{5} \delta_k NSI * FREQ\_k_{it} \qquad \text{Equations (3) and (3a)}$$

Variable	Predicted Sign	Estimation of Eq (3) (n = 151,532) *EA_RET* No Controls	Estimation of Eq (3) (n = 74,011) *EA_RET* With Controls[b]	Estimation of Eq (3a) (n = 151,498) *QRET* No Controls	Estimation of Eq (3a) (n = 74,112) *QRET* With Controls[b]
ΔAINC_0	+	0.233	0.313	0.884	1.284
		[0.0001][a]	[0.0001][a]	[0.0001]	[0.0001]
ΔAINC_1	+	0.149	0.239	0.573	0.861
		[0.0001]	[0.0001]	[0.0001]	[0.0001]
ΔAINC_2	+	0.128	0.193	0.637	0.933
		[0.0001]	[0.0001]	[0.0001]	[0.0001]
ΔAINC_3	+	0.122	0.179	0.524	0.790
		[0.0001]	[0.0001]	[0.0001]	[0.0001]
ΔAINC_4	+	0.075	0.189	0.321	0.740
		[0.0001]	[0.0001]	[0.0001]	[0.0001]
ΔAINC_5	+	0.099	0.195	0.371	0.956
		[0.0001]	[0.0001]	[0.0001]	[0.0001]
NSI_0_{it}	?	0.039	−0.130	0.158	0.278
		[0.0009]	[0.0358]	[0.0015]	[0.1026]
NSI_1_{it}	?	0.021	−0.006	0.117	0.317
		[0.0582]	[0.8948]	[0.0136]	[0.1928]
NSI_2_{it}	?	0.052	0.006	0.207	0.371
		[0.0001]	[0.9027]	[0.0001]	[0.1113]
NSI_3_{it}	?	0.036	0.007	0.221	0.357
		[0.0496]	[0.8902]	[0.0001]	[0.0454]

(continued on next page)

Panel A: Estimation of Equations (3) and (3a)

$$R_{it} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \phi_k \Delta AINC * FREQ\_k_{it} + \sum_{k=0}^{5} \delta_k NSI * FREQ\_k_{it} \qquad \text{Equations (3) and (3a)}$$

Variable	Predicted Sign	Estimation of Eq (3) (n = 151,532) *EA_RET* No Controls	Estimation of Eq (3) (n = 74,011) *EA_RET* With Controls[b]	Estimation of Eq (3a) (n = 151,498) *QRET* No Controls	Estimation of Eq (3a) (n = 74,112) *QRET* With Controls[b]
NSI_4_{it}	+	0.067	−0.039	0.372	0.538
		[0.0011]	[0.4881]	[0.0001]	[0.0053]
NSI_5_{it}	+	0.001	0.029	0.382	0.614
		[0.9632]	[0.5238]	[0.0001]	[0.0030]
Adj. R^2		0.0188	0.0268	0.0153	0.0418

Panel B: Difference Tests from Estimation of Equations (3) and (3a)

$$R_{it} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \phi_k \Delta AINC * FREQ\_k_{it} + \sum_{k=0}^{5} \delta_k NSI * FREQ\_k_{it} \qquad \text{Equations (3) and (3a)}$$

Variable	Predicted Sign	Estimation of Eq (3) *EA_RET* No Controls	Estimation of Eq (3) *EA_RET* With Controls	Estimation of Eq (3a) *QRET* No Controls	Estimation of Eq (3a) *QRET* With Controls
$\Delta AINC\_0 - NSI\_0$	$\varphi_0 > \delta_0$	0.193	0.444	0.726	1.006
		[0.0001][a]	[0.0001][a]	[0.0001][a]	[0.0001][a]
$\Delta AINC\_0 - NSI\_1$	?	0.212	0.319	0.766	0.968
		[0.0001]	[0.0001]	[0.0001]	[0.0002]
$\Delta AINC\_0 - NSI\_2$	?	0.181	0.307	0.677	0.913
		[0.0001]	[0.0001]	[0.0001]	[0.0001]
$\Delta AINC\_0 - NSI\_3$	?	0.196	0.306	0.662	0.927
		[0.0001]	[0.0001]	[0.0001]	[0.0001]

(continued on next page)

Panel B: Difference Tests from Estimation of Equations (3) and (3a)

$$R_{it} = \sum_{k=0}^{5} \beta_k FREQ\_k_{it} + \sum_{k=0}^{5} \phi_k \Delta AINC * FREQ\_k_{it} + \sum_{k=0}^{5} \delta_k NSI * FREQ\_k_{it} \qquad \textbf{Equations (3) and (3a)}$$

Variable	Predicted Sign	Estimation of Eq (3) *EA_RET* No Controls	Estimation of Eq (3) *EA_RET* With Controls	Estimation of Eq (3a) *QRET* No Controls	Estimation of Eq (3a) *QRET* With Controls
ΔAINC_0 – NSI_4	?	0.217	0.352	0.511	0.747
		[0.0001]	[0.0001]	[0.0001]	[0.0001]
ΔAINC_0 – NSI_5	?	0.232	0.284	0.502	0.670
		[0.0119]	[0.001]	[0.0001]	[0.0001]
NSI_5 – NSI_0	$\delta_5 > \delta_0$	−0.038	0.159	0.224	0.336
		[0.1194]	[0.0002]	[0.0070]	[0.0169]
NSI_5 – NSI_1	$\delta_5 > \delta_1$	−0.020	0.034	0.264	0.297
		[0.4118]	[0.1778]	[0.0025]	[0.1115]
NSI_5 – NSI_2	$\delta_5 > \delta_2$	−0.051	0.023	0.175	0.243
		[0.1423]	[0.3676]	[0.0554]	[0.1645]
NSI_5 – NSI_3	$\delta_5 > \delta_3$	−0.035	0.022	0.160	0.257
		[0.2150]	[0.4976]	[0.1065]	[0.0489]
NSI_5 – NSI_4	$\delta_5 > \delta_4$	−0.066	0.068	−0.002	0.076
		[0.0470]	[0.0673]	[0.9448]	[0.5697]

[a] Two-tailed p.

[b] For the tests reported in columns 4 and 6 we include a vector of controls variables that are included as main effects and interacted with each of the income and special item variables. See Appendix C for definitions of the control variables.

Variable Definitions:

R_{it} = $QRET_{it}$ or EA_RET_{it};

$QRET_{it}$ = quarterly market-adjusted stock return for firm i (adjusted for value-weighted index) accumulated over the period from one trading day after the earnings announcement date for the prior quarter $(t-1)$ through the earnings announcement date for the current quarter (t);

EA_RET_{it} = two-day earnings announcement period $[-1,0]$ market-adjusted stock return (adjusted for value-weighted index) for firm i in quarter t;

(continued on next page)

$RNSI_{it}$ = the amount of firm i's after-tax negative special item (Compustat data item #32). The negative special item is adjusted at the firm's quarterly effective tax rate in which the special item is recognized;
NSI_{it} = $RNSI_{it}$ scaled by the market value of equity in quarter t–4;
INC_{it-j} = income before extraordinary items (Compustat item #8) for firm i in quarter $t-j$, where j = 0 or 4;
$FREQ\_k_{it}$ = 1 if firm i reported k negative special items in the 12 quarters prior to quarter t, where k takes on values of 0 through 5, otherwise 0. When k takes on the value of 5, this indicates that firm i reported at least five negative special items in the 12 quarters prior to quarter t;
$\Delta AINC_{it}$ = unexpected earnings before NSIs for firm i in quarter t scaled by market value of equity in quarter t–4, measured as $(INC_{it} - RNSI_{it} - INC_{it-4}) / MVAL_{it-4}$, where $MVAL$ is market value of equity;
$\Delta AINC\_k_{it}$ = $\Delta AINC_{it} * FREQ\_k_{it}$;
NSI_k_{it} = $NSI * FREQ\_k_{it}$; and
$CONTROLS$ = See Appendix C for definitions of the control variables.

these variables lacks statistical significance at conventional levels ($p = 0.119$, see Panel B). Interestingly, when the control variables are introduced into the analysis, this pattern reverses as the *NSI_5* coefficient significantly exceeds the *NSI_0* coefficient ($p < 0.01$). That is, the marginal effect is that the market places greater weight on NSIs where there is a prior history of such charges in the short window around the earnings announcement.

Columns 5 and 6 report the estimation of Equation (3a), where we expand the return accumulation period to a quarterly window. Regardless of whether controls are included, these results reveal a declining pattern of $\Delta AINC\_k$ coefficients as frequency increases, consistent with the Equation (3) results. However, the *NSI_k* coefficient patterns from either estimation of Equation (3a) clearly differ from the results reported in the estimation of Equation (3) that does not include controls (i.e., the replication of Elliott and Hanna 1996). In either *QRET* estimation, the largest special item coefficient occurs for *NSI_5* (0.382 for the No Controls estimation and 0.614 for the estimation With Controls) and the second largest occurs for *NSI_4*. In Panel B, the coefficient on *NSI_5* exceeds that on *NSI_0* in either estimation of Equation (3a) (significance levels of 0.007 and 0.017, respectively). In spite of the significant increase in the coefficient on *NSI* with frequency, the $\Delta AINC\_0$ coefficient exceeds ($p < 0.01$) each of the *NSI_k* coefficients in every estimation. In addition to market returns adjusted for the value-weighted index, we estimate versions of Equations (3) and (3a) using size-adjusted and beta-adjusted returns with and without controls. The results of these additional tests, not tabulated, are quantitatively and qualitatively similar to the tabulated results. More importantly, our conclusions are robust to the use of these alternative return measures.[21]

By evaluating announcement period and quarterly return windows, we confirm the Elliott and Hanna (1996) conclusions with respect to earnings, but present uniquely different evidence with respect to the valuation of repetitive NSIs. The results from the quarterly return window clearly demonstrate a reduction in the earnings response coefficient associated with an increase in the reporting frequency of NSIs, which is the primary focus of their study. More important for our study, the quarterly return results as well as the earnings announcement period results with controls are clearly inconsistent with the notion that the valuation effects of NSIs decline with frequency. These results suggest the market prices repetitive NSIs as more persistent than nonrepetitive charges consistent with the NSI persistence tests we report in Tables 2 and 3.[22]

Interestingly, Table 4 reports that the Equation (3a) *NSI_k* coefficients converge to a magnitude roughly equal to the corresponding $\Delta AINC\_k$ coefficients once the number of prior charges exceeds three (particularly in the no-controls estimation). This raises the possibility that NSI persistence effects rise to the same level as the decreasing earnings persistence effects in such settings. On the other hand, the Equation (2) no-controls earnings persistence analysis reported in Table 3 does not support such an interpretation, as the *NSI_k* coefficients are all substantially smaller than the associated earnings coefficients. One likely explanation for at least part of this inconsistency is the substantial inter-period expense transfer effects documented by Burgstahler et al. (2002). They report approximately 8.7 and 27.7 percent of the quarter t NSI is expense transferred from quarters t+1 and t+4, respectively. If NSIs are used to accelerate future period expenses into a current quarter NSI, then this would result in a downward bias on the *NSI*

[21] In our primary tests, we measure prior reporting frequency based on the number of NSIs reported in the prior 12 quarters. Our primary empirical method ignores the possibility that the recurrence of *any* special item (positive or negative) provides information affecting the valuation of a current special item. We re-estimate Equation (3a) by redefining *FREQ* as the number of prior reported special items regardless of sign. The results of these additional tests, not tabulated, are quantitatively and qualitatively similar to the tabulated results.

[22] We ran positive special item versions of Equations (3) and (3a). The results reveal no evidence that the market places a higher multiple on repetitive positive special items.

coefficients in earnings persistence regressions. That is, in terms of earnings persistence, one dollar of future period expense accelerated into a current period NSI will, *ceteris paribus*, increase future earnings. Indeed, the significantly negative *NSI_0* coefficients observed in the no-controls estimations from Table 3 are highly consistent with the presence of such an inter-period transfer effect in our data. By contrast, the marginal persistence effect of a NSI is a decrease in future earnings since all of the *NSI_k* coefficients in the with-controls estimations in Table 3 are significantly positive. In addition, it is also conceivable that, since the market coefficients in Equation (3a) are estimates of the market's perception of the persistence of recurring NSIs relative to changes in earnings, the lack of symmetry between the market results from Equation (3a) and the actual persistence results from Equation (2) may reflect market overestimation of the persistence of these "recurring" charges.

Restructuring Charge Frequency Analysis

Table 5, Panel A reports regression results from the estimation of Equations (4) and (4a) which focus on restructuring charges rather than special items. The $\Delta AINC\_RC\_k$ coefficients for both equations loosely correspond to the patterns identified by Elliott and Hanna (1996) as well as in our own Table 4 analysis (i.e., a decline as frequency increases). Further, we find that when firms report three or more restructuring charges ($\Delta AINC\_RC\_3$ and $\Delta AINC\_RC\_4$), the earnings coefficients lack statistical significance. The short-window restructuring charge coefficients from Equation (4) generally lack statistical significance regardless of reporting frequency.[23] These results underscore: (1) the comparatively small sample of restructuring events in play here, and (2) the notion that prior disclosure of these events may be common.

The coefficients of central interest in our analysis, however, are the series of restructuring charge coefficient estimates in Equation (4a). With the expanded return accumulation window, we are more confident of capturing the value relevant information in restructuring charges. The coefficient on a current restructuring charge not preceded by a prior charge, *RC_0*, is −0.019 and insignificant. On the other hand, the coefficients on the second, third, and fourth or more restructuring charges (*RC_2*, *RC_3*, and *RC_4*) are all positive and significantly greater than zero ($p < 0.01$). In addition, the coefficients on *RC_2*, *RC_3*, and *RC_4* increase monotonically from 0.366 for *RC_2* to 0.637 for *RC_4*.

Table 5, Panel B reports tests of coefficient differences from the estimation of Equations (4) and (4a). Focusing on the results from the estimation of Equation (4a), we find that the $\Delta AINC\_RC\_0$ coefficient exceeds ($p < 0.10$ or better) the *RC_0*, *RC_1*, and *RC_2* coefficients. However, we find no statistical difference between $\Delta AINC\_RC\_0$ and the *RC_3* or *RC_4* coefficients. Also, the coefficient on *RC_4*, the restructuring with the longest sequence of prior reported restructuring charges, exceeds the *RC_0* and *RC_1* coefficients ($p < 0.01$), but not *RC_2* or *RC_3*. Similar to the results from the estimation of Equation (3a), the Equation (4a) estimates are inconsistent with the valuation effect of repetitive "nonrecurring items" declining with observed frequency.[24]

An important reason for employing quarterly return windows in our analysis is to address the possibility that NSIs and restructurings are announced prior to the earnings announcement date. In

[23] The sole exception is the coefficient on *RC_2*.

[24] In untabulated tests, we estimate Equations (4) and (4a) with the additional control variables identified as being significantly associated with NSI frequency. The results of these additional tests are consistent with our tabulated results in Table 5. In addition to market returns adjusted for the value-weighted index, we also estimate Equations (4) and (4a) using size-adjusted and beta-adjusted returns. The results of these additional tests, not tabulated, are quantitatively and qualitatively similar to the tabulated results and our conclusions are robust to the use of these alternative return measures.

TABLE 5

Cross-Sectional Regressions of Earnings Announcement and Quarterly Market-Adjusted Returns on Unexpected Earnings before Restructuring Charges and Restructuring Charges Conditioned on Reporting Frequency (n = 15,533)

Panel A: Estimation of Equations (4) and (4a)

$$R_{it} = \sum_{k=0}^{4} \beta_k FREQ\_k_{it} + \sum_{k=0}^{4} \theta_k \Delta AINC\_RC * FREQ\_k_{it} + \sum_{k=0}^{4} \gamma_k RC * FREQ\_k_{it} \qquad \text{Equations (4) and (4a)}$$

Variable	Predicted Sign	Estimation of Eq (4) Dependent Variable *EA_RET*	Estimation of Eq (4a) Dependent Variable *QRET*
ΔAINC_RC_0	+	0.292	0.863
		[0.0001][a]	[0.0001]
ΔAINC_RC_1	+	0.168	0.789
		[0.0136]	[0.0013]
ΔAINC_RC_2	+	0.133	0.704
		[0.0060]	[0.0514]
ΔAINC_RC_3	+	0.103	0.514
		[0.3654]	[0.2705]
ΔAINC_RC_4	+	0.055	0.305
		[0.6872]	[0.6620]
RC_0	+	0.012	−0.0193
		[0.9008]	[0.9720]
RC_1	+	0.030	0.111
		[0.3491]	[0.2754]
RC_2	+	0.127	0.366
		[0.0015]	[0.0082]
RC_3	+	0.084	0.446
		[0.1842]	[0.0017]
RC_4	+	0.032	0.637
		[0.4698]	[0.0009]
Adj. R^2		0.0243	0.0216

(continued on next page)

Panel B: Difference Tests from Estimation of Equations (4) and (4a)

$$R_{it} = \sum_{k=0}^{4} \beta_k FREQ\_k_{it} + \sum_{k=0}^{4} \theta_k \Delta AINC\_RC * FREQ\_k_{it} + \sum_{k=0}^{4} \gamma_k RC * FREQ\_k_{it} \qquad \textbf{Equations (4) and (4a)}$$

Variable	Predicted Sign	Estimation of Eq (4) With *EA_RET*	Estimation of Eq (4a) With *QRET*
$\Delta AINC\_RC\_0 - RC\_0$	$\theta_0 > \gamma_0$	0.280 [0.0001][a]	0.882 [0.0001][a]
$\Delta AINC\_RC\_0 - RC\_1$	?	0.227 [0.0086]	0.752 [0.0075]
$\Delta AINC\_RC\_0 - RC\_2$	?	0.165 [0.0146]	0.496 [0.0697]
$\Delta AINC\_RC\_0 - RC\_3$	?	0.209 [0.0541]	0.417 [0.3812]
$\Delta AINC\_RC\_0 - RC\_4$	?	0.261 [0.0098]	0.226 [0.9814]
$RC\_4 - RC\_0$	$\gamma_4 > \gamma_0$	0.020 [0.5551]	0.656 [0.0001]
$RC\_4 - RC\_1$	$\gamma_4 > \gamma_1$	0.002 [0.9619]	0.526 [0.0017]
$RC\_4 - RC\_2$	$\gamma_4 > \gamma_2$	−0.095 [0.0304]	0.271 [0.3670]
$RC\_4 - RC\_3$	$\gamma_4 > \gamma_3$	−0.052 [0.3514]	0.191 [0.8841]

[a] Two-tailed p.

Variable Definitions:

R_{it} = $QRET_{it}$ or EA_RET_{it};
$QRET_{it}$ = quarterly market-adjusted stock return for firm i (adjusted for value-weighted index) accumulated over the period from one trading day after the earnings announcement date for the prior quarter (t–1) through the earnings announcement date for the current quarter (t);
EA_RET_{it} = two-day earnings announcement period [−1,0] market-adjusted stock return (adjusted for value-weighted index) for firm i in quarter t;
RRC_{it} = the amount of firm i's after-tax restructuring charge, adjusted at the firm's quarterly effective tax rate in which the charge is recognized;
INC_{it-j} = income before extraordinary items (Compustat item #8) for firm i in quarter $t-j$, where j = 0 or 4;
$\Delta AINC\_RC_{it}$ = unexpected earnings before restructuring charges for firm i in quarter t measured as $(INC_{it} - RC_{it} - INC_{it-4}) / MVAL_{it-4}$, where *MVAL* is market value of equity;

(continued on next page)

RC_{it} = RRC_{it} scaled by the market value of equity in quarter t–4;
$FREQ\_k_{it}$ = 1 if firm i reported k restructuring charges in the 12 quarters prior to quarter t, where k takes on values of 0 through 4. When k takes on the value of 4, this indicates that firm i reported at least four restructuring charges in the 12 quarters prior to quarter t;
$\Delta AINC\_RC\_k_{it}$ = $\Delta AINC\_RC_{it} * FREQ\_k_{it}$, and
RC_k_{it} = $RC_{it} * FREQ\_k_{it}$.

fact, we find that over 65 percent of the restructurings in our sample are announced four or more days prior to the earnings announcement date. If the two-day earnings announcement period return window utilized by Elliott and Hanna (1996) and replicated by us in Equations (3) and (4) results in downward biased *NSI* (*RC*) coefficients as we suggest, then a restructuring announcement return window should yield results qualitatively similar to our quarterly return results from the estimation of Equations (3a) and (4a). Accordingly, we estimate the following modified version of Equation (4):

$$RC\_RET_{it} = \sum_{k=0}^{4} \beta_k FREQ\_k_{it} + \sum_{k=0}^{4} \gamma_k RC{*}FREQ\_k_{it}. \quad (5)$$

RC_RET_{it} is the two-day restructuring announcement period [−1,0] market-adjusted stock return for firm i in quarter t, and all other variables are as defined previously. Equation (5) excludes earnings variables because this model focuses on the market return around the announcement of the restructuring. Restructuring announcement dates are identified by searching Lexis-Nexis for the public announcement of the restructuring for the 1,220 observations in our sample. If no separate announcement date was found, then we assume the announcement date for the restructuring is the earnings announcement date.

Table 6 reports results from estimating Equation (5). The coefficient on a current restructuring charge not preceded by a prior charge, *RC_0*, is 0.001 and insignificant. The coefficients on *RC_1* and *RC_2* are also insignificant. On the other hand, and in contrast to the earnings announcement return test results reported in Table 5 (Equation (4)), the coefficients on *RC_3* and *RC_4* are positive and significantly greater than 0 ($p < 0.10$). In addition, rather than declining or flattening, the *RC_k* coefficients increase monotonically from 0.001 for *RC_0* to 0.217 for *RC_4*. Table 6 also reports tests of coefficient differences from the estimation of Equation (5). Here we find that the coefficient on *RC_4* is significantly greater than the coefficients on *RC_0*, *RC_1*, and *RC_2* ($p <$ 0.05 or better), but there is no significant difference between *RC_4* and *RC_3*. Overall, the results from the estimation of Equation (5) suggest that the valuation effect of repetitive "nonrecurring items" increases with observed frequency. In addition, these results are consistent with the notion that use of an earnings announcement date return window biases coefficients on restructuring (*NSI*) variables toward zero.[25]

Supplemental Analyses

Tax Rate Adjustment

In our tabulated empirical tests, we adjust special items for the firm-specific quarterly effective tax rate. To assess the sensitivity of our results to our choice of the tax rate, we perform several additional analyses utilizing alternative tax rate adjustments. First, we estimate an implied tax rate on NSIs using the following empirical model:

$$TaxExp_{it} = \psi_0 + \psi_1 ADJ\_PTI_{it} + \psi_2 PSI_{it} + \psi_3 NSI_{it} + \psi_4 PSI * P\_FREQ_{it} + \psi_5 NSI * N\_FREQ_{it} + \psi_6 ADJ\_PTI * P\_FREQ_{it} + \psi_7 ADJ\_PTI{*}N\_FREQ_{it}. \quad (6)$$

INC_{it} is income before extraordinary items (Compustat data item #8) for firm i in quarter t; PTI_{it} is pre-tax income (Compustat data item #23) for firm i in quarter t; RSI_{it} is special items (Compustat data item #32) reported by firm i in quarter t; *PSI* equals *RSI* if $RSI > 0$, otherwise $PSI = 0$; *NSI* equals *RSI* if $RSI < 0$, otherwise $NSI = 0$; ADJ_PTI_{it} is PTI_{it} less RSI_{it}; N_FREQ_{it}

[25] In untabulated tests, we re-estimate Equation (5) with the additional control variables as before. The results of this additional test are consistent with our tabulated results in Table 6.

TABLE 6

Cross-Sectional Regressions of Two-Day Restructuring Announcement Period Market-Adjusted Returns on Restructuring Charges Conditioned on Reporting Frequency (n = 1,220)

$$RC\_RET_{it} = \sum_{k=0}^{4} \beta_k FREQ\_k_{it} + \sum_{k=0}^{4} \gamma_k RC * FREQ\_k_{it} \qquad \textbf{Equation (5)}$$

Variable	Predicted Sign	Coefficient Estimates two-tailed p]
RC_0	?	0.001 [0.9802][a]
RC_1	?	0.010 [0.7826]
RC_2	?	0.034 [0.3662]
RC_3	+	0.091 [0.0643]
RC_4	+	0.217 [0.0006]
RC_4 – RC_0	$\gamma_4 > \gamma_0$	0.216 [0.0036]
RC_4 – RC_1	$\gamma_4 > \gamma_1$	0.207 [0.0043]
RC_4 – RC_2	$\gamma_4 > \gamma_2$	0.183 [0.0128]
RC_4 – RC_3	$\gamma_4 > \gamma_3$	0.126 [0.1485]
Adj. R^2		

[a] Two-tailed p.

Variable Definitions:

RC_RET_{it} = two-day restructuring announcement period [−1,0] market-adjusted stock return (adjusted for value-weighted index) for firm i in quarter t;

RRC_{it} = the amount of firm i's after-tax restructuring charge, adjusted at the firm's quarterly effective tax rate in which the charge is recognized;

RC_{it} = RRC_{it} scaled by the market value of equity in quarter t–4;

$FREQ\_k_{it}$ = 1 if firm i reported k restructuring charges in the 12 quarters prior to quarter t, where k takes on values of 0 through 4. When k takes on the value of 4, this indicates that firm i reported at least four restructuring charges in the 12 quarters prior to quarter t; and

$RC\_k_{it} = RC_{it} * FREQ\_k_{it}$.

(P_FREQ_{it}) is the number of NSIs (PSIs) reported by firm i over the 12 quarters prior to quarter t; IMP_TE_{it} is the implied tax expense measured as PTI_{it} less INC_{it}; $ACT\_TE_{,it}$ is actual tax expense (Compustat data item #6) for firm i in quarter t; and $TaxExp_{it}$ is ACT_TE_t or IMP_TE_t. All continuous variables are scaled by the market value of equity in quarter t–4.

The empirical evidence from estimating Equation (6), not tabulated, is consistent with prior research that finds positive and negative special item observations are taxed at substantially dif-

ferent rates. Specifically, Beaver et al. (2007) report evidence consistent with the conclusion that negative special item observations have effective tax rates close to zero, while positive special item observations have effective tax rates that approach the highest marginal statutory tax rates. We find that the tax rate on negative special items implied by the coefficient on NSI ranges between 4 and 6 percent while the tax rate on positive special items is approximately 19 percent. In addition, the estimated tax rate declines as the prior reporting frequency of NSIs increases (i.e., significantly negative coefficient on the $NSI * N\_FREQ_{it}$ interaction). From this analysis, we incorporate an alternative tax rate adjustment. Specifically, we adjust positive special items at the average of the coefficient on *PSI*, 19 percent. For NSIs, the tax adjustment is based on the prior reporting frequency of NSIs. That is, we incorporate the specific tax rate conditioned on prior reporting frequency. In addition to this alternative adjustment, we also incorporate two additional tax rate adjustments. First, we adjust negative and positive special items using firm-specific annual marginal tax rates provided by Graham (1996a, 1996b). Second, we adjust all special items at the highest statutory rate, 35 percent for our sample time period. We re-estimate all of our empirical analyses using each of these alternative tax rate adjustments and the results from these alternative specifications are robust and our conclusions are unaffected.

Expectation of Special Items

Our primary valuation models measure unexpected earnings as the difference between the earnings in the current quarter and earnings in quarter t–4. On the other hand, we make no seasonal or other adjustments to NSIs. This approach is consistent with the view that investors base expectations of continuing earnings on prior observations, but, in general, do not expect a special item to be recognized. There are two reasons why this approach may not be valid. First, it seems unlikely that the same earnings news specification would be appropriate for both the short-window and long-window analysis. Second, the underlying premise of our analysis is that NSIs are likely to recur for some companies. If so, it seems reasonable that under certain circumstances investors will expect nonzero NSIs and the design should recognize this expected NSI effect. That is, the amount of "news" in a negative special item should depend on whether the firm recognizes special items frequently.[26]

To address these issues, we ran two additional specifications of our models. First, we ran a levels version of Equation (3a) using current quarter earnings adjusted for the NSI rather than seasonally differenced earnings. The results from this additional test are consistent with our tabulated results. Specifically, we find that as the prior sequence of NSIs increases, the coefficient on *NSI* increases. Second, we added a new "unexpected NSI" analysis. In this additional analysis, we measure expected NSIs as the average NSI over the prior 12 quarters and the unexpected NSI as the reported NSI less the expected component. We then partition NSIs into two components (expected and unexpected) in Equation (3a) and rerun our analyses. With respect to the unexpected NSI tests, it appears that our tabulated results are primarily attributable to the unexpected component of NSIs. In particular, we find that the coefficients on unexpected NSIs increase as the prior reporting frequency increases, while the coefficient on the expected component is generally insignificant regardless of the prior reporting history.

Cross-Sectional Determinants of ERCs

The longer return window estimation inferences come at a cost. The quarterly return accumulation windows employed in Equations (3a) and (4a) raise the possibility that risk-related factors may confound the analysis. Accordingly, we modify Equations (3a) and (4a) to include

[26] We thank an anonymous reviewer for bringing both of these issues to our attention.

controls for cross-sectional variables shown by prior research to affect the relation between returns and unexpected earnings (e.g., Kormendi and Lipe 1987; Collins and Kothari 1989; Easton and Zmijewski 1989). The controls we use are growth (market-to-book ratio), risk (beta),[27] leverage (long-term debt divided by the sum of long-term debt plus common stockholders' equity), and firm size (log of total assets). The results from this analysis are essentially identical to those reported in Tables 4 and 5 and do not affect our conclusions.[28]

VI. CONCLUSION

This study examines the persistence and market implications of what has been referred to in the business media as "recurring nonrecurring" items (Stewart 2002). Our results suggest that the frequency of prior reported NSIs is positively related to both the frequency and magnitude of subsequently reported NSIs. Our results also suggest that the frequency of prior reported NSIs has a significant effect on the persistence of earnings before NSIs. Overall, our evidence suggests that as the sequence of prior reported NSIs increases, the future earnings implications of a current NSI become more predictable.

A prior related analysis by Elliott and Hanna (1996) finds that the market places greater weight on initial NSIs than subsequent NSIs (in the two-day window around the earnings announcement). That is, such reoccurrences are viewed as decreasing in value relevance by market participants. Our analysis confirms this basic finding for NSIs in the short window around the earnings announcement date. This result is perplexing, given the clear relation between increased frequency of NSIs and the likelihood and magnitude of future NSI recognition. However, when we introduce control variables for NSI likelihood or expand the return accumulation window to allow for pre-announcement disclosure of such events, we observe a valuation effect opposite that reported in Elliott and Hanna (1996). Specifically, we find that as prior NSI frequency increases, the weight the market attaches to such subsequent charges increases rather than decreases. Moreover, this increase in valuation is consistent with the estimated persistence properties we document for such high-frequency charges.

Interestingly, the decline in the valuation of operating earnings with special items frequency that is of central interest in Elliott and Hanna (1996) holds for the expanded return window. The robustness of this decline suggests that it is unlikely that one of the three explanations they identify—heightened pre-announcement of earnings information by large write-off firms (Elliott and Hanna 1996, 153)—is the underlying source of this phenomenon. That is, this study's expanded return accumulation period includes the time period over which pertinent pre-disclosure earnings information is released, thereby capturing these information effects. Hence, our empirical method is not sensitive to the specific date that the market learns of the nonrecurring item, as long as it occurs within the quarter-long accumulation window we employ.

Overall, our empirical findings are important in at least two respects. First, our results suggest that NSIs may be indicative of the firm's operating environment and nearly as repetitive as other components of income. Alternatively, repetitive write-offs may be reflective of a reporting strategy that classifies ordinary expenses as NSIs to manipulate pre-write-off income (i.e., "classification

[27] The firm-specific beta is calculated using the 200 trading days prior to the beginning of quarter t with the CRSP equally weighted portfolio as the market proxy.

[28] In addition to the pooled cross-sectional return test results that we tabulate, we use three alternative methods to calculate t-statistics in our return tests. First, we re-estimate all return regressions on an annual basis using the Fama and MacBeth (1973) methodology. We then estimate mean coefficients and t-statistics from these annual regressions. Second, we calculate robust-standard errors using two-way clustering by firm and year to control for both time-series and cross-sectional dependence (Gow et al. 2010). Third, we calculate robust-standard errors using two-way clustering by industry and year to control for the industry clustering we note in Appendix B. Our findings are robust to these alternative methods.

shifting" as McVay 2006; Fan et al. 2010). Either interpretation suggests the objective of separately reporting an item below the ordinary operating elements of the income statement to highlight its unusual or infrequent nature is being compromised. Second, our results suggest media and regulator concern about the use of repetitive NSIs to fool the market may be unfounded. Specifically, our results suggest that the market values these "recurring nonrecurring" items more like the other components of recurring earnings.

Our analysis provides limited insights about manager motivation to engage in repetitive reporting of NSIs. The inclusion of control variables (including interactions with NSI amounts) for factors associated with NSI reporting frequency generally yields stronger rather than weaker results in our analyses. To the extent that this set of control variables is effective in controlling for such factors, the results suggest that managerial discretion regarding NSI reporting is an underlying driver of the persistence and valuation effects we observe. Moreover, given that our analysis also suggests the market seems to understand and unwind the repetitive nature of "nonrecurring" charges, the question arises as to why managers continue to record such charges. The provocative answer to this question is that managers are irrepressibly and irrationally opportunistic in reporting earnings. Perhaps managers believe that market participants value such charges as transitory, even when they become recurring. Indeed, anecdotal evidence (e.g., Lalli and Lim 1997) supports this notion. A more mundane alternative is that managers are sometimes slower than the market to respond to changing conditions. Hence, managers continue to report items as nonrecurring based on industry conventions or experience, failing to recognize that the basic operating environment has changed in such a way that existing financial reporting conventions are no longer appropriate. That is, what once was nonrecurring or rare has now become ordinary and expected and is valued accordingly.

APPENDIX A

Type, Number, Frequency, and Repetitive Nature of Negative Special Items[a]

Type of Special Item	Number of Observations Current Quarter	Percent of Current Quarter Observations	Number of the Same Type *SI* over the Prior 12 Quarters	Mean Frequency of the Same Type *SI* over the Prior 12 Quarters[b]	Number of the Different Type *SI* over the Prior 12 Quarters	Mean Frequency of the Different Type *SI* over the Prior 12 Quarters[c]
Restructuring Charges	54	34	124	2.30	321	27.8%
Merger-Related Costs	6	4	33	5.50	36	47.8%
In-Process R&D	11	7	29	2.64	45	39.2%
Asset Write-offs	19	12	25	1.32	136	15.5%
Litigation Costs	8	5	10	1.25	45	18.2%
Goodwill Write-off	36	23	103	2.86	239	30.1%
Asset Disposition	8	5	8	1.00	38	17.4%
Other	17	11	30	1.76	104	22.4%

[a] This table was created by directly examining the income statements and footnote disclosures for the current and prior 12 periods for 100 randomly selected firm-quarters characterized by at least three large special items (greater than 1 percent of assets) in the prior 12 quarters. The number of current quarter observations exceeds 100 because some firms reported multiple types of special items in a single quarter. From the footnote disclosures, we classified the current NSIs (*SI*) into one of the eight tabulated types based on the description of the charge in the footnote. The number over the prior 12 quarters represents a total of all of the items found in the prior 12 quarters, not just those over 1 percent of assets.

[b] This number is taken by dividing the total number of observations in each category from the prior 12 quarters by the number of observations in the category. For example, there were 124 restructuring charges reported in the prior 12 quarters by the 54 firms that reported a restructuring charge in the sample quarter (124/54 = 2.30).

[c] This percentage is the number of same type *SI* over the prior 12 quarter divided by the sum of the number of same and different type *SI*s.

APPENDIX B

Mean and Median Number of Prior Negative Special Items by Industry Group

Fama-French Industry Classification	n[a]	Mean[b]	Median	Lower Quartile	Upper Quartile
Agriculture	65	2.55	2	1	4
Food Products	308	1.29	1	0	2
Candy and Soda	33	0.79	0	0	1
Alcoholic Beverages	66	1.58	1	0	3
Tobacco Products	29	1.66	1	1	2
Recreational Products	159	1.58	1	0	2
Entertainment	344	1.40	1	0	2
Printing and Publishing	212	1.87	1	1	2
Consumer Goods	407	1.97	1	0	2
Apparel	312	1.67	1	0	2
Healthcare	427	1.87	1	0	3
Medical Equipment	727	1.90	1	1	3

(continued on next page)

Fama-French Industry Classification	n[a]	Mean[b]	Median	Lower Quartile	Upper Quartile
Pharmaceutical Products	1,120	1.89	1	1	3
Chemicals	439	1.79	1	1	3
Rubber and Plastic Products	185	1.79	1	0	3
Textiles	131	1.51	1	0	2
Construction Materials	349	1.70	1	0	3
Construction	191	1.36	1	0	2
Steel Works, etc.	302	1.60	1	0	2
Fabricated Products	112	1.20	1	0	2
Machinery	785	1.75	1	1	3
Electrical Equipment	232	1.65	1	0	2
Miscellaneous	236	1.77	1	0	2
Automobiles and Trucks	220	1.30	1	0	2
Aircraft	99	1.53	1	0	2
Shipbuilding, Railroad Equip	40	1.67	1	0	3
Defense	21	1.24	1	1	2
Precious Metals	93	1.39	1	0	2
Nonmetallic Mining	52	0.67	0	0	1
Coal	11	1.18	1	1	1
Petroleum and Natural Gas	536	1.04	1	0	2
Utilities	496	0.96	1	0	1
Telecommunications	767	1.34	1	0	2
Personal Services	181	1.50	1	0	2
Business Services	3,868	2.33	2	1	3
Computers	1,468	2.36	2	1	3
Electronic Equipment	1,761	2.11	2	1	3
Measuring and Control Equip	598	1.77	1	1	
Business Supplies	405	1.57	1	0	2
Shipping Containers	89	1.44	1	1	2
Transportation	343	1.19	1	0	2
Wholesale	729	1.52	1	0	2
Retail	963	1.33	1	0	2
Restaurants, Hotel, Motel	514	1.51	1	0	2
Banking	1,187	0.24	0	0	0
Insurance	526	0.84	0	0	1
Real Estate	114	1.58	1	0	2
Trading	803	0.84	0	0	1

[a] n = number of quarterly observations with a negative special item.

[b] This represents the mean number of NSIs reported over the prior 12 quarters conditioned on the recognition of a current quarter NSI.

(continued on next page)

APPENDIX C

Control Variables

Variable	Name	Data Source	Definition
Accounts Receivable Turnover	*AR_TURN*	Compustat	Accounts receivable turnover measured as sales revenue (Compustat data item #12) divided by receivables−total (Compustat data item #2) for firm *i* measured in the year prior to the year in which quarter *t* falls.
Inventory Turnover	*INV_TURN*	Compustat	Inventory turnover measured as cost of goods sold (Compustat data item #41) divided by the sum of inventories−total (Compustat data item #3) and inventory LIFO reserve (Compustat data item #240) for firm *i* measured in the year prior to the year in which quarter *t* falls.
PP&E Turnover	*PPE_TURN*	Compustat	PP&E turnover measured as sales revenue (Compustat data item #12) divided by PP&E−total, net (Compustat data item #8) for firm *i* measured in the year prior to the year in which quarter *t* falls.
Sales Revenue per Employee	*SALE_EMP*	Compustat	Sales revenue per employee measured as sales revenue (Compustat data item #12) divided by the number of employees (Compustat data item #29) for firm *i* measured in the year prior to the year in which quarter *t* falls.
Profit Margin	*PM*	Compustat	Profit margin measured as income before extraordinary items (Compustat data item #18) divided by sales revenue (Compustat data item #12) for firm *i* measured in the year prior to the year in which quarter *t* falls.
Change in ROE	ΔROE	Compustat	Change in return on equity measured as the change income before extraordinary items (Compustat data item #18) scaled by shareholder's equity (Compustat data item #216) for firm *i* measured in the year prior to the year in which quarter *t* falls.

(continued on next page)

Variable	Name	Data Source	Definition
Change in ROA	*ΔROA*	Compustat	The four-year average change in return on assets measured as income before extraordinary items (Compustat data item #18) scaled by total assets (Compustat data item #6) for firm *i* measured over the four years prior to the year in which quarter *t* falls
Growth in Industry Sales	*IND_GROWTH*	Compustat	The four-year average percentage growth in industry sales revenue (Compustat data item #12) for firm *i* measured over the four years prior to the year in which quarter *t* falls.
One Year Return	*RET1*	CRSP	One-year raw market return for firm *i* measured over the year prior to the beginning of quarter *t*.
Five Year Return	*RET5*	CRSP	Four-year raw market return for firm *i* measured from year −5 through the end of the year prior to the beginning of quarter *t*.
Book-to-Market Ratio	*BTM*	Compustat	Book-to-market ratio measured as common equity (Compustat data item #60) divided by the market value of equity (Compustat data item #199 times item #25) for firm *i* measured in the year prior to the year in which quarter *t* falls.
Change in the Book-to-Market Ratio	*ΔBTM*	Compustat	The four-year average change in the BTM for firm *i* measured over the four years prior to the year in which quarter *t* falls.
Change in the Industry BTM Ratio	*IND_ΔBTM*	Compustat	The four-year average change in the industry BTM for firm *i* measured over the four years prior to the year in which quarter *t* falls.
Firm Size	*SIZE*	Compustat	Natural log of total assets (Compustat data item #6) for firm *i* measured at the end of quarter *t*−1.
Prior Year Loss Indicator	*LOSS*	Compustat	Indicator variable equal to 1 if firm had a loss (Compustat data item #18 less than 0) for any of the three years prior to the year in which quarter *t* falls, otherwise 0.

REFERENCES

Accounting Principles Board (APB). 1973. *Reporting the Results of Operations—Reporting the Effects of Disposal of a Segment of A Business, and Extraordinary, Unusual and Infrequently Occurring Events and Transactions.* Opinion No. 30. New York, NY: AICPA.

Adut, D., W. Cready, and T. J. Lopez. 2003. Restructuring charges and CEO cash compensation: A reexamination. *The Accounting Review* 78 (Jan): 169–192.

Beaver, W., M. McNichols, and K. Nelson. 2007. An alternative interpretation of the discontinuity in earnings distributions. *Review of Accounting Studies* 12 (December): 525–556.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Source of Collinearity*. New York, NY: John Wiley.

Bens, D. A., and R. Johnston. 2009. Accounting discretion: Use or abuse? An analysis of restructuring charges surrounding regulator action. *Contemporary Accounting Research* 26 (Fall): 673–699.

Bleakley, F. 1995. New write-offs mostly please investors. *Wall Street Journal* (December 21): C1.

Bradshaw, M., and R. Sloan. 2002. GAAP versus the street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (March): 41–66.

Bricker, R., G. Previts, T. Robinson, and S. Young. 1995. Financial analyst assessment of company earnings quality. *Journal of Accounting, Auditing & Finance* (Summer): 541–554.

Burgstahler, D., J. Jiambalvo, and T. Shevlin. 2002. Do stock prices fully reflect the implications of special items for future earnings? *Journal of Accounting Research* 40 (June): 585–612.

Carter, M. E. 2000. Operating performance following corporate restructurings. Working paper, Columbia University.

Chaney, P., C. Hogan, and D. Jeter. 1999. The effect of reporting restructuring charges on analysts' forecast revisions and errors. *Journal of Accounting and Economics* 27 (July): 261–284.

Clement, M. B., L. Koonce, and T. J. Lopez. 2007. The roles of task-specific forecasting experience and innate ability in understanding analyst forecasting performance. *Journal of Accounting and Economics* 44 (December): 378–398.

Collins, D. W., and S. P. Kothari. 1989. An analysis of the inter-temporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics* 11 (July): 143–181.

DeAngelo, H., L. DeAngelo, and D. Skinner. 1994. Accounting choice in troubled companies. *Journal of Accounting and Economics* 17 (January): 113–143.

Dechow, P., M. Huson, and R. Sloan. 1994. The effect of restructuring charges on executives' cash compensation. *The Accounting Review* 69 (January): 138–156.

Dial, J., and K. Murphy. 1995. Incentives, downsizing, and value creation at General Dynamics. *Journal of Financial Economics* 27: 261–314.

DiMartino, D. 2002. Earnings confusion persists amid increased frequency of restructuring charges. *The Dallas Morning News* (November 4).

Easton, P., and M. Zmijewski. 1989. Cross-sectional variation in the stock market response to the announcement of accounting earnings. *Journal of Accounting and Economics* 11 (July): 117–142.

Elliott, J., and J. D. Hanna. 1996. Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research* 34 (Supplement): 135–155.

———, and W. Shaw. 1988. Write-offs as accounting procedures to manage perceptions. *Journal of Accounting Research* 26 (Supplement): 91–119.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *The Journal of Political Economy* 81 (May-June): 607–636.

Fan, Y., A. Barua, W. Cready, and W. Thomas. 2010. Managing earnings using classification shifting: Evidence from quarterly special items. *The Accounting Review* 85 (4): 1303–1323.

Financial Accounting Standards Board and International Accounting Standards Board (FASB and IASB). 2008. *Financial Statement Presentation—Joint Project of the IASB and FASB*. Norwalk, CT: FASB.

Fowler, T. 2006. The Enron trial: A few pennies can buy deception. *Houston Chronicle* (February 5): A1.

Francis, J., J. D. Hanna, and L. Vincent. 1996. Causes and effects of discretionary asset write-offs. *Journal of Accounting Research* 34 (Supplement): 117–134.

Freeman, R., and S. Tse. 1992. A nonlinear model of security price response to unexpected earnings. *Journal of Accounting Research* 30 (Autumn): 185–209.

Gow, I., G. Ormazabal, and D. Taylor. 2010. Correcting for the cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (March): 483–512.

Graham, J. R. 1996a. Debt and the marginal tax rate. *Journal of Financial Economics* 41 (May): 41–73.

——— 1996b. Proxies for the corporate marginal tax rate. *Journal of Financial Economics* 42 (October):

187–221.

Kormendi, R., and R. Lipe. 1987. Earnings innovations, earnings persistence, and stock returns. *The Journal of Business* 60 (July): 323–345.

Kothari, S., and R. Sloan. 1992. Information in prices about future earnings: Implications for earnings response coefficients. *Journal of Accounting and Economics* 15 (June-September): 143–171.

Lalli, F., and P. Lim. 1997. The real Dow. *Money* (March): 76–79.

Lipe, R. C. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (Supplement): 37–64.

Lopez, T. 2002. Evidence on the incremental information contained in the components of restructuring charges. *Journal of Business Finance & Accounting* 29 (June): 613–659.

McVay, S. 2006. Earnings management using classification shifting: An examination of core earnings and special items. *The Accounting Review* 81 (May): 501–532.

Parkash, M. 1995. The differential information contents of unexpected permanent and temporary earnings. *Journal of Business Finance & Accounting* 22 (July): 695–713.

Soffer, L. C., S. R. Thiagarajan, and B. Walther. 2000. Earnings preannouncement strategies. *Review of Accounting Studies* 5 (March): 5–26.

Smith, R., and S. Lipin. 1996. Are companies using restructuring costs to fudge the figures? A repeated strategic move makes future earnings seem unrealistically rosy. *Wall Street Journal* (January 30): A1.

Stewart, J. 2002. Nonrecurring charges keep recurring in company financial reports. *Chicago Tribune* (March 24): 1.

Strong, N., and M. Walker. 1993. The explanatory power of earnings for stock returns. *The Accounting Review* 68 (April): 385–399.

Ziegler, B. 1996. IBM will account for job reduction costs as cost of business, not as write-down. *Wall Street Journal* (March 6): C1.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1617–1646

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1617

Analysts' Incentives to Overweight Management Guidance When Revising Their Short-Term Earnings Forecasts

Mei Feng
University of Pittsburgh

Sarah McVay
The University of Utah

ABSTRACT: We document that, when revising their short-term earnings forecasts in response to management guidance, analysts wishing to curry favor with management weight the guidance more heavily than predicted, based on the credibility and usefulness of the guidance. This overweighting of guidance is present prior to equity offerings and other events that could lead to investment banking business. Although analysts sacrifice their forecast accuracy by overweighting management guidance, they appear to benefit, on average, by subsequently gaining the underwriting business for their banks. Thus, while analysts wishing to please managers are optimistic in their long-term earnings forecasts, they take their cue from management when determining their short-term earnings forecasts.

Keywords: *analyst incentives; management guidance; management forecasts; analyst forecasts; analyst revisions.*

Data Availability: *The data used in this study are publicly available from the sources indicated in the text.*

I. INTRODUCTION

Analysts face many competing incentives: they not only want to be accurate to signal their skill and build their reputation, but they also want to issue optimistic long-term forecasts to please the firms they follow or to generate trading volume for their brokerage houses (Mikhail et al. 1999; Lim 2001; Hong and Kubik 2003; Cowen et al. 2006; Libby et al. 2008). Optimism is not necessarily desirable, however, for short-term forecasts (Matsumoto 2002; Richardson et al. 2004; Ke and Yu 2006). Thus, we expect analysts with incentives to please the firms

We thank two anonymous reviewers, Bonita Austin, Brian Cadman, Asher Curtis, Harry Evans, Zhaoyang Gu, Chan Li, Arianna Pinello, Steve Kachelmeier (the editor), Scott Schaefer, Jenny Tucker, and workshop participants at the 2007 American Accounting Association Annual Meeting, the 2008 Financial Accounting and Reporting Section Midyear Meeting, New York University, the University of Pittsburgh, and The University of Utah for their helpful comments and suggestions.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: February 2008
Accepted: December 2009
Published Online: August 2010

they follow to take their cue from management when determining their short-term forecasts. Specifically, when revising their forecasts in response to management guidance, we expect that these analysts will weight the guidance more heavily than predicted based on the perceived credibility and usefulness of the guidance (hereafter, overweight management guidance).[1] Although by overweighting management guidance analysts bear the cost of reduced forecast accuracy, this behavior pleases management, as it helps align the market's expectations with those of management (e.g., Givoly and Lakonishok 1979; Ajinkya and Gift 1984; King et al. 1990; Stickel 1991).

We examine analysts' incorporation of guidance prior to equity offerings, where analysts have been shown to have strong incentives to curry favor with managers. The equity offering setting is advantageous because the benefit to analysts is straightforward (underwriting fees for their banks) and we can identify which investment banks underwrote the deal, *ex post*. In addition, we can conduct powerful cross-sectional tests by analyzing different analysts' incorporation of the same guidance, as we expect only those analysts with incentives to please management to overweight management guidance.

We find that analysts overweight management guidance for firms that subsequently announce an equity offering. After controlling for other variables that affect analysts' incorporation of guidance, such as the credibility and usefulness of the forecast, analysts on average incorporate 27 percent more of the management guidance into their revised short-term forecasts for firms that announce an equity offering in the subsequent year. This overweighting is economically significant, as, on average, analysts tend to incorporate approximately 50–80 percent of management guidance (e.g., Baginski and Hassell 1990; see current Table 4).

Although we control for a variety of determinants related to the credibility and usefulness of management guidance, it is possible that guidance prior to equity offerings is informative in ways not captured by our measures of usefulness and credibility, leading to the appearance of an overweighting of guidance. To address this possibility, we conduct two sets of tests to better distinguish between the expected incorporation of credible and useful guidance and the overweighting of guidance to please managers prior to equity offerings.

First, we examine analysts' *ex post* forecast accuracy. The incorporation of credible and useful guidance generally improves analysts' earnings forecasts. Thus, if the overweighting of guidance we observe is in response to some omitted credibility and usefulness measures, then analysts' forecasts should be more accurate as a result. In contrast, we find that analysts' *ex post* forecast errors increase by 40 percent, on average, as a result of the overweighting of management guidance prior to equity offerings. This finding is consistent with analysts bearing significant accuracy costs in order to please management.

Second, we investigate whether the weight placed on management guidance prior to equity offerings varies with analyst independence and affiliation. These cross-sectional tests use only management guidance issued by firms that announce an equity offering in the next year, thereby mitigating the endogeneity concerns associated with the equity offering. We hold variation in the credibility and usefulness of the guidance constant by requiring that both analyst types (e.g., affiliated and unaffiliated) forecast earnings for a given firm, allowing us to determine how *different analysts* respond to the *same guidance*.

Our first cross-sectional test examines differences between dependent and independent analysts. Only analysts employed by investment banks (dependent analysts) have the potential to gain

[1] To clarify, we expect analysts to revise their forecasts either upward or downward depending on management's preference, which is communicated via management guidance. We document that analysts overweight guidance in both directions in Section V.

the underwriting business for their firms. Thus, if the overweighting of guidance stems from analysts' incentives to gain underwriting business, we should find that dependent analysts, but not independent analysts, overweight management guidance prior to equity offerings. Conversely, if the underlying properties of firms issuing equity cause the overweighting, we should not observe a systematic difference by analyst type. We find that dependent analysts weight the same management guidance more heavily than independent analysts, consistent with the overweighting of guidance representing a mechanism used by analysts to curry favor with management.

Our second cross-sectional test examines differences between affiliated and unaffiliated analysts. This test allows us to condition on the *ex post* realization of gaining the underwriting fees. If managers reward analysts who overweight their guidance, we should find that those analysts employed by the banks subsequently underwriting the offerings (affiliated analysts) weight the firm's guidance more heavily than other analysts. Results are consistent with this conjecture; we find that analysts' overweighting of management guidance is associated with obtaining the underwriting business of the upcoming equity offering.[2]

We conduct several additional analyses to corroborate these findings. First, to further control for the endogeneity of an equity offering, we model the equity issuance choice and conduct a two-stage Heckman analysis. We continue to find evidence that analysts overweight management guidance prior to equity offerings. We also assess the likelihood that the inclusion of an omitted variable could overturn the results, following Frank (2000), and conclude that the probability is small.[3] Second, we document that analysts who overweight management guidance also tend to issue strong buy recommendations, a previously documented mechanism to please management (e.g., Michaely and Womack 1999), consistent with analysts simultaneously employing several tools to curry favor with management. Third, we illustrate that our conclusions do not require analysts' explicit knowledge of an equity offering. We find that analysts overweight management guidance for firms *expected* to have equity offerings. Moreover, the overweighting of guidance is more prevalent in the six months prior to the equity offering, relative to offerings announced 7–12 months in the future, consistent with analysts being more responsive when the potential payoffs are nearer (and presumably more certain). Fourth, we document that the overweighting of guidance prior to announcements of equity offerings is present for both upward and downward guidance, consistent with analysts following management in either direction in order to curry favor. Finally, we also find some evidence that analysts overweight management guidance prior to other events that provide analysts an incentive to please management: i.e., prior to debt offerings and M&A activity, where debt underwriting fees and M&A advisory fees might be gained by the analyst's employer.

To summarize, we find that analysts overweight short-term management guidance when they have an incentive to please management—preceding an equity offering. This overweighting does not appear to be a function of analysts incorporating useful information, but rather results in less accurate forecasts. The overweighting of guidance is, on average, present only when the analysts are able to gain the underwriting business and the analysts appear to benefit from this behavior by increasing their chances of gaining the subsequent underwriting business for their banks.

---

[2] An alternative explanation for this result is that the subsequent underwriter is also the past underwriter, and the analysts are attempting to please management to *maintain* the underwriting business. For example, Michaely and Womack (1999) find that brokerage analysts are more optimistic in their recommendations for companies their firm has recently taken public. This plausible alternative does not alter our main conclusion. Analysts under either scenario overweight management guidance in order to please management. Moreover, Krigman et al. (2001) find that approximately one-third of lead underwriters change from deal to deal; thus, the past underwriters cannot automatically assume they will be the subsequent underwriters, further heightening their incentive to please management.

[3] Though we explore this endogeneity concern in detail in Section V, we cannot completely rule out the possibility that the endogenous nature of equity offerings drives our results.

Our study contributes to the understanding of how incentives to please management affect analysts' short-term earnings forecasts. Although analysts wishing to please managers are overly optimistic in their long-term earnings forecasts and recommendations, we document that they take their cue from management when determining their short-term earnings forecasts. The desire to please management, therefore, can result in systematic errors in analysts' short-term earnings forecasts. Although we use the setting of equity offerings to document analysts' overweighting of guidance to please management, the implications of our findings are much broader. This mechanism can be used by any analyst wishing to curry favor with management, conditional on the issuance of management guidance, whether the analyst wishes to gain bank fees (the focus of the current study), to receive preferred treatment in conference calls (Libby et al. 2008; Mayew 2008), or to please management for other reasons (e.g., Chen and Matsumoto 2006; Ke and Yu 2006).

II. MOTIVATION OF HYPOTHESIS

In general, analysts have an incentive to issue accurate earnings forecasts. Hong and Kubik (2003) find that the least accurate analysts are relegated to less desirable brokerage houses, while more accurate analysts are more likely to move up to high-status brokerage houses. Lim (2001) concludes that a positive forecast bias is the outcome of rational analysts trying to minimize their forecast error while also trying to please management. Consistent with this reasoning, Chan et al. (2003a) find that analysts consistently exceed realized growth when estimating long-term growth in earnings, and Dechow et al. (2000) find that the most optimistic long-term growth forecasts are for equity issuers.[4] This positive bias also extends to stock recommendations. For example, Dugar and Nathan (1995) document that analysts employed by investment banks issue more optimistic stock recommendations than do analysts employed by non-investment banks, and Lin and McNichols (1998) find that those analysts subsequently employed as the underwriters (affiliated analysts) issue more optimistic recommendations and long-term growth forecasts than do unaffiliated analysts.

Although optimistic recommendations and long-term forecasts are arguably Pareto-optimal, this is not the case for short-term forecasts. Blanket optimism for short-term earnings forecasts is costly if the firm misses the benchmark (e.g., Skinner and Sloan 2002) and, consistent with this, prior research has not found systematic evidence of optimism in short-term forecasts (e.g., Lin and McNichols 1998). There is no clear strategy for analysts to garner management's favor with their short-term forecasts. Some managers will be pleased by optimistic forecasts, for example, if they plan to issue equity or sell their shares before earnings are reported, if they are in danger of losing their positions, or if the firm is in danger of a debt-rating downgrade or bankruptcy. Other managers would prefer pessimistic forecasts, for instance if they face high litigation risk, if their bonuses depend on meeting the analyst forecast, or if they plan to issue equity or sell their shares after earnings are reported and want to positively surprise the market to maximize the sale price (Rogers and Stocken 2005).

In the face of these varying incentives, managers can use their guidance to move expectations to the desired levels (e.g., Ajinkya and Gift 1984; King et al. 1990; Hirst et al. 2008). Thus, rather than observing systematic optimism or pessimism in shorter-term analyst forecasts, we expect that, to curry favor with managers, analysts will take their cue from management and overweight management guidance when revising their short-term earnings forecasts, even though the overweighting will likely result in less accurate analyst forecasts.

---

[4] It is also possible that analysts tend to follow firms that they expect to perform well and choose not to follow firms they expect to perform poorly, leading to the appearance of *ex post* optimism (Hayes 1998; Bradshaw et al. 2006).

Several conditions must exist for analysts to overweight guidance to please managers. First, managers must want analysts' forecasts to closely mirror their own guidance. This desire on the part of managers is reasonable because, when analysts incorporate guidance, it aids in the alignment of market expectations with those of management (e.g., Givoly and Lakonishok 1979; Lys and Sohn 1990; Stickel 1991). We therefore assume that managers will desire analysts to be responsive to their guidance.

Second, managers must be able to identify and reward analysts who overweight their guidance. We assume that a manager can infer each analyst's weighting, relative to an unbiased weighting of the guidance, based on the manager's knowledge of his or her own reputation, forecast credibility, and forecast precision. Managers can reward the overweighting by giving the analyst (or associated investment bank) special consideration, for example, in the underwriter-selection process.[5] Finally, analysts must have an incentive to curry favor with management. We use upcoming equity offerings to proxy for analysts' incentives to please management. During the time period examined, underwriting departments were known to award significant bonuses to analysts who help attract new underwriting business (Cowen et al. 2006, 122). We expect analysts employed by an investment bank to have incentives to please management in this setting in order to gain the underwriting business for their firm.

In sum, we expect analysts wishing to please management to overweight management guidance when revising their short-term earnings forecasts, at the expense of their *ex post* accuracy, in order to benefit from management's special consideration—for example, by increasing the likelihood of gaining underwriting business for their investment bank.

Hypothesis: Analysts overweight management guidance when revising their short-term earnings forecasts, at the expense of their *ex post* forecast accuracy, to curry favor with management.

We conduct several analyses that jointly test our hypothesis. First, we examine whether analysts overweight management guidance when revising their short-term earnings forecasts for firms announcing an equity offering in the next year. Second, we examine whether this overweighting of guidance leads to larger analyst forecast errors. Third, we conduct cross-sectional tests to determine whether only analysts who could feasibly gain the underwriting business overweight management guidance and if this behavior appears to result in gaining the underwriting business.

III. DATA AND DESCRIPTIVE STATISTICS

Data and Sample

We use data from 1994–2005, obtaining quarterly management guidance, analyst forecasts, and actual (reported) earnings from the First Call database. We acquire information regarding the timing and underwriter of the equity offering from SDC Platinum. We obtain the necessary data for the control variables from the 2005 Quarterly and Annual Compustat File. We outline the sample selection procedure in Table 1. As Figure 1 illustrates, we consider management guidance for quarter q that was issued after the earnings announcement for quarter $q-1$ and before the earnings announcement for quarter q. We examine the analyst forecast revisions for quarter q around the corresponding management guidance. In this short-term setting, analysts cannot be systematically optimistic in their forecasts as some managers will prefer pessimistic forecasts. Thus, analysts must take their cue from management's guidance when determining what short-

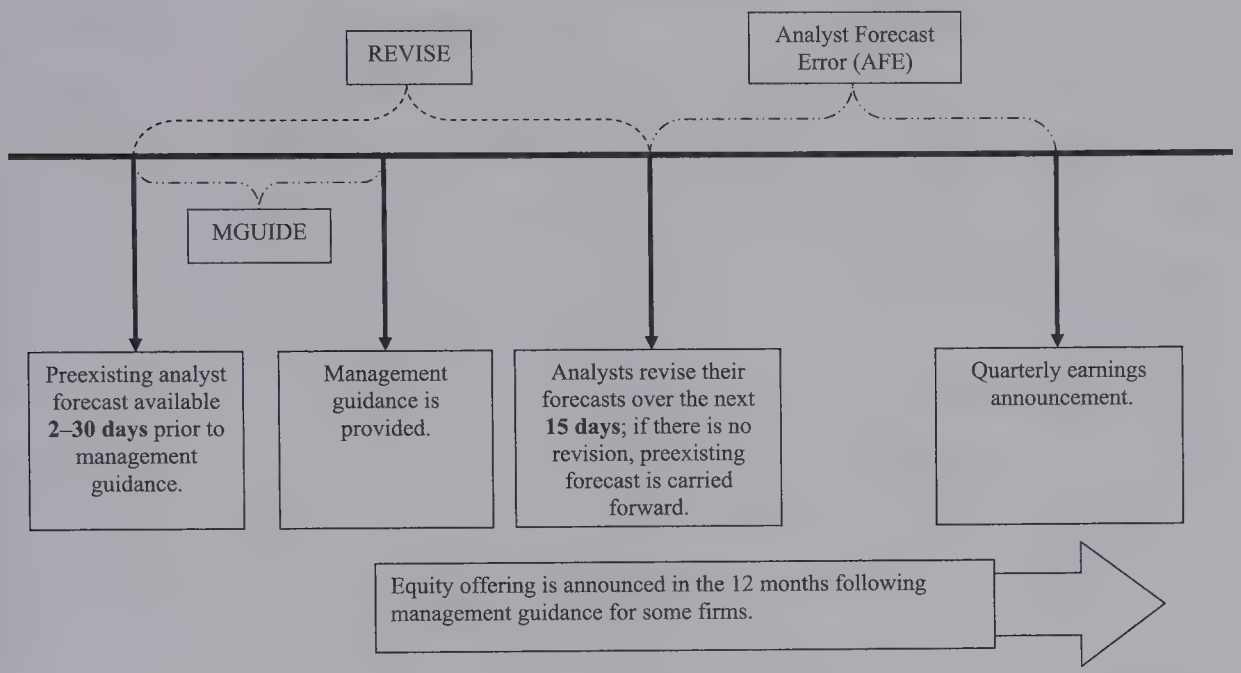
[5] Although our focus is on underwriting fees, this special consideration could also include other types of consulting fees, inside information, and conference call favoritism. Clearly, managers will also consider deal terms, underwriter quality, and other criteria when determining whom to select as underwriters. These additional features of the choice of underwriter should add noise to the tests, thereby biasing against finding evidence consistent with our hypothesis.

TABLE 1

Management Guidance Sample Selection 1994–2005

	Number of Forecasts	Number of Firms
Total Forecasts	43,885	5,983
Less:		
Non-EPS forecasts	1,291	137
Non-USD currency forecasts	295	132
Forecasts unduly affected by mergers or accounting changes	169	21
Forecasts issued before 1994	58	13
EPS forecasts made after 1994	42,072	5,680
Less:		
Qualitative forecasts	8,278	911
Point and range forecasts	33,794	4,769
Less:		
Duplicate forecasts	33	1
Forecasts missing actual earnings data	4,131	203
Forecasts with actual earnings data	29,630	4,565
Less:		
Forecasts without CUSIPs	60	32
Forecasts with missing Compustat earnings announcement dates	4,089	370
Forecasts issued before the earnings announcement date of the prior quarter	12,655	358
Forecasts issued after the earnings announcement date of the current quarter	236	44
Final management guidance sample	12,590	3,761
Less:		
Observations without analysts' preexisting forecasts	5,912	1,445
Management guidance with analysts' forecasts	6,678	2,316
Less:		
Observations without stock price data	974	554
Management guidance with stock price as a scalar	5,704	1,762
Less:		
Multiple guidance issuances (keeping the most recent guidance)	515	0
The top and bottom one percent of management guidance, analysts' revisions, and managerial reputation	204	54
Final Sample	4,985	1,708

FIGURE 1
Timeline and Variable Measurement



term earnings forecast will curry favor.[6] If a firm provides multiple forecasts for the same quarter, then we keep the most recent (last-issued) guidance.[7] The final sample contains 4,985 firm-quarter observations consisting of 1,708 unique firms.[8]

Variable Definitions and Descriptive Statistics

All variables used in the analysis are defined in Table 2, with corresponding descriptive statistics in Table 3. Our research design requires that we measure management's "suggested" revision (based on their public guidance) and analysts' subsequent forecast revisions to their short-term earnings forecast. The higher the correlation between management and analyst revisions, the greater the "weighting" of guidance by analysts. In the following section, we introduce a number of control variables that are expected to affect the weight placed on the management guidance, thereby allowing us to focus on the "overweighting" of guidance by analysts prior to equity offerings.

Management's suggested revision (*MGUIDE*) is the measure of how far the management

[6] We include pre-announcements in the sample because, in our setting, it is not clear that analysts' incentives would change within the window leading up to the earnings announcement. Specifically, we expect analysts to wish to please management throughout the quarter, including during the "pre-announcement" period. Our results do not change qualitatively when we exclude pre-announcements from our sample (not tabulated).

[7] We delete 515 instances of subsequently revised management guidance for a given firm-quarter, keeping only the most recently issued guidance in a given firm-quarter. It is possible that firms issuing multiple forecasts in a given quarter are systematically different from the rest of the sample. We find no statistical difference in analysts' incorporation of guidance for these 515 firm-quarters in the multivariate tests (not tabulated).

[8] After the sample constraints, we retain approximately 29 percent of the unique firms from the initial population. We compare the size, performance, and expected growth of the final sample to the initial population from which the sample was collected. We find that our sample is larger, more profitable, and has higher expected growth, consistent with analysts following these types of firms (Bhushan 1989; McNichols and O'Brien 1997).

TABLE 2
Variable Definitions

Variable	Description	Calculation (Compustat data item numbers in parentheses)
SEO	If the firm issued equity	An indicator variable that is set equal to 1 if the firm announces an equity offering in the year following the management guidance, and 0 otherwise.
MGUIDE	The magnitude of the revision suggested by the manager	The difference of the management earnings forecast and the preexisting median consensus analyst forecast, scaled by stock price at the beginning of the quarter. The preexisting consensus analyst forecast is the most recent consensus before the management guidance (within 2–30 days). The midpoint of the range is used for all management guidance that is provided as a range forecast.
REVISE	The magnitude of the analyst forecast revision	The difference of the revised consensus analyst forecast and the preexisting consensus analyst forecast, scaled by stock price at the beginning of the quarter. The preexisting consensus analyst forecast is the most recent consensus before the management guidance (within 2–30 days). The revised consensus analyst forecast is the updated consensus forecast issued within 15 days of the management guidance. If there is not a revised consensus analyst forecast, then *REVISE* is 0.
FREQUENCY	The number of management forecasts	The number of quarterly management forecasts issued over the previous three years.
EPS	Earnings per share	Earnings per share scaled by the stock price at the beginning of the quarter.
|*AFE*|	The absolute *ex post* analyst forecast error	The absolute value of the median consensus analyst forecast error 15 days following management guidance, scaled by the stock price at the beginning of the quarter.
AFFILIATED	Affiliated analyst	An indicator variable that is set equal to 1 if the analyst is employed by the investment bank that subsequently underwrites the equity offering, and 0 otherwise.
DEPENDENT	Analyst employed by investment bank	An indicator variable that is set equal to 1 if the analyst is employed by an investment bank, and 0 otherwise.
DOWN	Downward management guidance	An indicator variable that is set equal to 1 if the management guidance falls below the preexisting consensus analyst forecast, and 0 otherwise.
REPUTATION	Historical reputation of the manager	The average accuracy of the preceding management guidance issued in the prior three years, where the accuracy is equal to 1, 0 and −1, respectively, if the absolute value of the preexisting analyst forecast error is greater than, equal to, or less than the absolute value of the management forecast error.
AGREE	Congruent signs on management guidance and market response	An indicator variable that is set equal to 1 if the three-day abnormal (market-adjusted) return around the management guidance (−1, +1) has the same sign as the management revision, and 0 otherwise. The abnormal return is equal to the difference between the firm return and the value-weighted market return.

(continued on next page)

TABLE 2 (continued)

Variable	Description	Calculation (Compustat data item numbers in parentheses)
HORIZON	Management guidance horizon	The log of the number of days between the issuance of management guidance and the earnings announcement.
RANGE	Management guidance precision	The range of the forecast scaled by the stock price at the beginning of the quarter for a range forecast, and 0 for a point forecast.
ANALYSTS	Analyst following	The log of the number of analyst forecasts used to form the consensus analyst forecast following management guidance.
$\lvert\Delta EPS\rvert$	Earnings shocks	The absolute change in earnings per share from First Call, scaled by the stock price at the beginning of the quarter.
LOSS	Loss indicator	An indicator variable that is set equal to 1 if earnings per share from First Call is negative, and 0 otherwise.
SPECIAL ITEMS	Special items	Special items from Compustat (#32), scaled by the market value of common stock at the beginning of the quarter.
FIRM SIZE	Firm size	The log of the market value of common stock at the beginning of the quarter.
ANALYST DISPERSION	Analyst disagreement	The standard deviation of the individual analyst forecasts used to generate the consensus analyst forecast following management guidance, scaled by the stock price at the beginning of the quarter.
DEBT	If the firm issued debt	An indicator variable that is set equal to 1 if the firm announces a debt offering in the year following the management guidance, and 0 otherwise.
M&A	If the firm had an M&A	An indicator variable that is set equal to 1 if the firm announces a merger or acquisition in the year following the management guidance, and 0 otherwise.

earnings forecast deviates from the preliminary earnings expectations: the management earnings forecast less the preexisting median consensus analyst forecast (the last consensus available in the 2–30 days prior to the guidance), scaled by price at the beginning of the quarter. The analyst revision (*REVISE*) is the revised consensus analyst forecast (the last revised consensus available in the 15 days following the management guidance), less the preexisting median consensus analyst forecast, scaled by price at the beginning of the quarter. In other words, *REVISE* is how much the consensus analyst forecast moves after management guidance is issued.[9] The magnitude of the mean analyst revision (*REVISE*; −0.0029) falls just below that of the "suggested" management

[9] We consider the consensus analyst forecast, rather than each individual analyst revision, to provide direct evidence that the consensus forecasts can be moved by the overweighting of guidance. We consider the median, rather than the mean, consensus analyst forecast because the mean consensus could be more easily moved by one analyst's overweighting of guidance to please management and, thus, examining the mean forecast could overstate the economic significance of our findings. Finally, we examine analyst revisions in the 15 days following management guidance to better capture revisions made by analysts in response to management guidance, rather than other news.

revision (*MGUIDE*; −0.0031), on average (see Table 3).

We measure equity offerings (*SEO*) as an indicator variable that equals 1 if the firm announces an equity offering in the year following the issuance of the management guidance, and equals 0 otherwise. We consider the announcement date instead of the actual offering date for two reasons. First, analysts cannot revise their forecasts following the announcement of equity offerings due to SEC-mandated quiet periods (e.g., see Lin and McNichols 1998).[10] Second, firms have usually selected their underwriters by the time of the announcement. Thus, it is before the announcement that analysts will be trying to gain the underwriting business. Approximately 4 percent of the sample issues equity through a secondary equity offering in the year following management guidance (Table 3).[11]

We compare guidance that is issued in the year prior to an equity offering announcement (SEO) to guidance issued during all other times (non-SEO) in Table 3. The frequency of guidance is higher among non-SEO observations relative to SEO observations, and the absolute value of analyst forecast errors is greater among SEO observations, though we explicitly link this error to the overweighting of guidance in the following section. Interestingly, we find no difference in the propensity of losses, the magnitude of special items, firm size, or analyst dispersion between SEO and non-SEO observations; thus, the results do not suggest that the larger forecast errors are a result of greater uncertainty prior to SEOs. Overall, there is no systematic evidence of one of these two groups clearly dominating the other with respect to the usefulness or credibility of guidance. Non-SEO observations tend to have more downward guidance (which is deemed more credible), while guidance among the SEO observations tends to be made by managers with slightly higher reputations.

IV. TEST DESIGN AND RESULTS

Do Analysts Overweight Management Guidance Prior to Equity Offerings?

To determine if analysts wishing to please management overweight management guidance when revising their short-term earnings forecasts, we estimate the following regression equation:

$$\begin{aligned} REVISE = {} & \alpha_0 + \alpha_1 MGUIDE + \alpha_2 MGUIDE \times SEO + \alpha_3 MGUIDE \times DOWN \\ & + \alpha_4 MGUIDE \times REPUTATION + \alpha_5 MGUIDE \times AGREE + \alpha_6 MGUIDE \\ & \times HORIZON + \alpha_7 MGUIDE \times RANGE + \alpha_8 MGUIDE \times ANALYSTS + \alpha_9 SEO \\ & + \alpha_{10} DOWN + \alpha_{11} REPUTATION + \alpha_{12} AGREE + \alpha_{13} HORIZON + \alpha_{14} RANGE \\ & + \alpha_{15} ANALYSTS + \varepsilon. \end{aligned} \quad (1)$$

The results are robust to several alternative specifications. First, our results do not change qualitatively when we use the mean consensus analyst forecasts instead of the median (not tabulated). Second, instead of analyst forecast revisions in 15 days, we examine the revisions in 30 days and up to the day before the earnings announcement. Our conclusions remain unchanged (not tabulated). Third, we include all analysts in the revised consensus figure; as an additional robustness check, we drop from this calculation analysts who did not revise their forecast; results are slightly stronger after the exclusion of these analysts (not tabulated). Finally, we also consider controlling for the age of the preexisting forecast, which is between two and 30 days. The age of the preexisting consensus is not a significant explanatory variable and does not change the results (not tabulated).

[10] Consistent with the interactions between managers and analysts preceding the announcement date, in our sample we find only 38 instances where management guidance is issued after the announcement of the equity offering but before 25 days following the actual offering (not tabulated).

[11] As a robustness check, we follow Bradshaw et al. (2006) and consider a second measure of equity financing that incorporates a broader set of equity financing (18 percent of the sample). Results are similar using this alternative financing measure (not tabulated).

TABLE 3
Descriptive Statistics

	Full Sample						SEO Observations (202 max. obs.)		Non-SEO Observations (4,783 max. obs.)	
Variable	n	Mean	Std. Dev.	Q1	Median	Q3	Mean	Median	Mean	Median
SEO	4,985	0.0405	0.1972	0.0000	0.0000	0.0000	1.0000	1.0000	0.0000***	0.0000***
MGUIDE	4,985	−0.0031	0.0072	−0.0044	−0.0006	0.0000	−0.0021	0.0000	−0.0031**	−0.0007***
REVISE	4,985	−0.0029	0.0066	−0.0042	−0.0003	0.0000	−0.0018	0.0000	−0.0029**	−0.0004***
EPS/PRICE	4,985	0.0093	0.0184	0.0041	0.0103	0.0168	0.0080	0.0098	0.0093	0.0103
FREQUENCY	4,985	1.6752	2.2472	0.0000	1.0000	2.0000	0.7228	0.0000	1.7155***	1.0000***
|AFE|	4,937	0.0012	0.0021	0.0000	0.0005	0.0013	0.0016	0.0008	0.0012	0.0005***
DOWN	4,985	0.5585	0.4966	0.0000	1.0000	1.0000	0.4505	0.0000	0.5630***	1.0000***
REPUTATION	3,026	0.5480	0.5431	0.0000	0.7500	1.0000	0.6261	1.0000	0.5456	0.7143**
AGREE	4,985	0.6088	0.4881	0.0000	1.0000	1.0000	0.5594	1.0000	0.6109	1.0000
HORIZON	4,985	32.5470	22.0381	15.0000	28.0000	43.0000	31.2822	27.0000	32.6005	28.0000
RANGE	4,985	0.0014	0.0023	0.0000	0.0008	0.0018	0.0015	0.0006	0.0014	0.0008**
ANALYSTS	4,937	9.3352	6.2863	5.0000	8.0000	13.0000	7.5931	7.0000	9.4103	8.0000***
|ΔEPS|	3,863	0.0100	0.0185	0.0018	0.0044	0.0111	0.0102	0.0054	0.0100	0.0044
LOSS	4,937	0.1485	0.3556	0.0000	0.0000	0.0000	0.1667	0.0000	0.1477	0.0000
SPECIAL ITEMS	4,937	−0.0084	0.0309	−0.0010	0.0000	0.0000	−0.0088	0.0000	−0.0084	0.0000
FIRM SIZE	4,937	6.9987	1.7280	5.7271	6.8597	8.1428	6.8927	6.7308	7.0033	6.8656
ANALYST DISP.	4,937	0.0011	0.0025	0.0002	0.0004	0.0010	0.0012	0.0006	0.0011	0.0004
DEBT	4,985	0.0628	0.2426	0.0000	0.0000	0.0000	0.1980	0.0000	0.0571***	0.0000
M&A	4,985	0.0401	0.1963	0.0000	0.0000	0.0000	0.0198	0.0000	0.0410	0.0000

*, **, *** Indicates that the mean and median are significantly different between Non-SEO and SEO observations at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test.

There are a maximum of 4,985 firm-quarter observations from 1994–2005; some variables have fewer observations due to missing data. The untransformed variables are presented for both *HORIZON* and *ANALYSTS* for ease of interpretation; the logged values are included in the following regression analyses.

Variables are defined in Table 2.

In general, analysts place weight on (or incorporate) management guidance when revising their forecasts; thus, without control variables, we expect α_1 to be positive and significant. To determine if analysts *overweight* guidance prior to equity offerings, we must first control for expected variation in the incorporation of guidance. We expect analysts to weight guidance more heavily if it is useful and credible (Hassell et al. 1988; Baginski and Hassell 1990; Williams 1996). We interact six measures of usefulness and credibility with *MGUIDE* to control for the expected incremental incorporation of guidance. The main effects of these variables are expected to be largely unassociated with analysts' revisions; thus, we include them for completeness but do not tabulate them. Because a firm can appear multiple times in the regression, we cluster the standard errors by firm following Petersen (2009) in this and all subsequent tests. All reported p-values are two-tailed.

We first control for two measures of credibility. *DOWN* is an indicator variable that equals 1 if the manager's suggested revision lowers expectations, which is generally viewed as more credible (Hassell et al. 1988; Cotter et al. 2006). *REPUTATION* is the average management guidance accuracy relative to analyst forecasts over the preceding three years (Williams 1996; Hutton and Stocken 2009). A higher value indicates that prior management guidance was more credible—i.e., had analysts incorporated the guidance, they would have reported a lower forecast error than if they had retained their original forecast.[12] We expect analysts to weight the guidance more heavily when it is more credible.

The next three controls proxy for the usefulness of the guidance. The first, *AGREE*, is an indicator variable that equals 1 if the three-day abnormal return around the management guidance has the same sign as the revision implied by the management guidance; this reflects the market's assessment of the usefulness of the guidance (e.g., Baginski and Hassell 1990). The second, *HORIZON*, is the log of the number of days prior to the earnings announcement that the guidance was issued. We expect guidance issued earlier in the quarter to be more useful to analysts. Finally, we include a measure of the precision of the guidance (*RANGE*), where we expect more precise guidance to be more useful (Baginski et al. 1993). We expect analysts to weight useful guidance more heavily when revising their short-term earnings forecasts. Because the measures in the analysis are based on the consensus analyst forecast, we also control for the log of the number of analysts (*ANALYSTS*) following the firm.[13]

We predict that, prior to equity offerings, analysts will overweight management guidance when revising their short-term earnings forecasts. Therefore, we expect α_2 to be positive and significant.

Table 4 presents the results from estimating Regression (1). Referring to the first column of results, which does not include control variables and thus includes all 202 firms with subsequent announcements of equity offerings, α_1 is positive and significant (0.83; $p < 0.01$), consistent with analysts incorporating a significant portion of management guidance into their short-term forecasts. This magnitude is significantly higher than that documented in the early work by Baginski

[12] Bhojraj et al. (2008) find that managers who are frequent forecasters also tend to be more accurate, and analysts tend to incorporate more of their guidance. Because our measure of reputation does not take forecasting frequency into account, as a sensitivity check we create the variable *#FORECASTS* that is equal to the number of forecasts issued by the firm over the three-year period that *REPUTATION* is measured. Consistent with Bhojraj et al. (2008), we find that the interaction term on *MGUIDE* × *REPUTATION* × *#FORECASTS* is positive and significant ($\alpha = 0.0422$; $p < 0.01$). Our inferences do not change; the coefficient on the interaction of *MGUIDE* and *SEO* is 0.2765 ($p < 0.01$), which is similar to the main results, presented in Table 4.

[13] We do not have a predicted sign on the interaction term of *ANALYSTS* and *MGUIDE*. On one hand, because more analysts would need to revise their forecasts to move the consensus when there are more analysts following the firm, it is possible that the guidance is incorporated less when there are more analysts. On the other hand, analyst following is not exogenous, and it is likely that more analysts follow larger and relatively more important firms, which might suggest a greater degree of incorporation of management guidance.

TABLE 4

Revision Magnitude on Equity Offerings

		Dependent Variable = *REVISE*	
Independent Variables	**Predicted Sign**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
Intercept		−0.0003***	−0.0012*
		(−6.77)	(−1.80)
MGUIDE	+	0.8303***	−0.1347
		(36.43)	(−0.31)
MGUIDE × *SEO*	+	0.1053***	0.2718***
		(3.14)	(3.68)
MGUIDE × *DOWN*	+		0.1473**
			(1.98)
MGUIDE × *REPUTATION*	+		0.1442***
			(2.92)
MGUIDE × *AGREE*	+		0.1585***
			(2.96)
MGUIDE × *HORIZON*	+		0.1005
			(0.99)
MGUIDE × *RANGE*	−		−7.7507**
			(−2.53)
MGUIDE × *ANALYSTS*	?		0.1151***
			(2.88)
Main Effects		Included	Included
Number of SEO Observations		202	87
Total Number of Observations		4,985	3,025
Adjusted R^2		0.8204	0.8803

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test. The regressions use all available observations. See Table 1 for sample selection and Table 2 for data definitions. All t-statistics are based on robust standard errors clustered at the firm level.

and Hassell (1990). Recently, however, management guidance has become more precise, more accurate, and is often supplemented with additional information, consistent with guidance issued during our time period being more credible and useful (Baginski et al. 2008). Consistent with our expectations, the coefficient on the interaction between management guidance and equity offerings, α_2, is positive and significant (0.11; $p < 0.01$), indicating that analysts place a greater weight on management guidance when the firm subsequently announces an equity offering.

In the next column we interact each of the six control variables with management guidance, as management guidance may be more useful or credible prior to equity offerings.[14] The main effect of *MGUIDE* becomes insignificant, consistent with the bulk of the analyst revision being explained by the credibility and usefulness of the management guidance. In initial support of our

[14] The sample is significantly reduced at 43 (61) percent of the original sample for SEO (all) firm-year observations. This reduction is almost entirely attributable to the requirement of *REPUTATION*, where firms offering equity tend to have issued less guidance in the past (see Table 3) and thus are missing this variable more often. Results are similar if we exclude this control variable from the analysis (not tabulated).

hypothesis, α_2 remains positive and significant (0.27; $p < 0.01$). In economic terms, analysts incorporate 27 percent more of the management guidance into their short-term earnings forecasts prior to equity offerings. This overweighting is economically significant, as, on average, analysts tend to incorporate approximately 80 percent of management guidance.

The coefficients on the control variables are consistent with our expectations. Analysts incorporate more of the management guidance when it is downward or issued by a more reputable manager (when it is more credible), as well as when the market reaction is in the same direction as the guidance and when the precision of the guidance is higher (when it is more useful). Finally, analysts tend to weight the guidance more heavily when more analysts follow the firm. In sum, while analysts weight guidance more heavily when it is credible and useful, they also place greater weight on guidance received prior to the announcement of an equity offering. Thus, results are consistent with our hypothesis that analysts with incentives to please management overweight management guidance when revising their short-term earnings forecasts.

Do Analysts Sacrifice Accuracy to Overweight Guidance Prior to Equity Offerings?

It is possible that management guidance prior to equity offerings is more informative in ways not captured by our measures of usefulness and credibility, leading to the appearance of an overweighting of guidance. To help distinguish between the expected incorporation of credible and useful guidance (Hassell et al. 1988; Baginski and Hassell 1990; Williams 1996; Cotter et al. 2006) and the overweighting of guidance associated with attempting to please managers prior to equity offerings (our hypothesis), we examine *ex post* analyst forecast errors.

If analysts weight guidance more heavily because it is more informative before equity offerings, or because analysts face more uncertainty regarding the growth or profitability of firms about to issue equity, then the greater reliance on management guidance should reduce analyst forecast errors. Alternatively, as suggested by Lim (2001) and Hong and Kubik (2003), when competing incentives are introduced, analysts are willing to sacrifice accuracy.[15] We predict that analysts overweight management guidance in order to increase their chances of gaining the underwriting business and, thus, this increased reliance on management guidance should reduce forecast accuracy.

In the following set of tests, we explicitly link the overweighting of guidance prior to equity offerings (*SEO REVISION* below) with larger analyst forecast errors in a multivariate estimation (Regression (2) below). First, we generate an estimate of the incremental weight placed on guidance prior to equity offerings from Regression (1) by decomposing the total analyst revision into three components:

(1) that associated with the upcoming equity offering:

$$SEO\ REVISION = \hat{\alpha}_2 MGUIDE \times SEO + \hat{\alpha}_9 SEO;$$

(2) that associated with other known factors affecting analyst responsiveness:

$$STANDARD\ REVISION = \hat{\alpha}_0 + \hat{\alpha}_1 MGUIDE + \hat{\alpha}_3 MGUIDE \times DOWN + \hat{\alpha}_4 MGUIDE \times REPUTATION + \hat{\alpha}_5 MGUIDE \times AGREE + \hat{\alpha}_6 MGUIDE \times HORIZON + \hat{\alpha}_7 MGUIDE \times RANGE + \hat{\alpha}_8 MGUIDE$$

[15] Ke and Yu (2006) examine the intertemporal characteristics of analysts' earnings forecasts, documenting that analysts switching from initially optimistic forecasts to *ex post* pessimistic forecasts just before the earnings announcement are more likely to receive inside information from managers (prior to Regulation Fair Disclosure) and more likely to be retained by their brokerage houses. Similarly, in an experimental setting, Libby et al. (2008) find that analysts are more likely to switch from initially optimistic forecasts to *ex post* pessimistic forecasts to gain favored conference call participation and information access. These analysts *ex post* have lower analyst forecast errors (as they gain inside information) and, thus, this setting differs from ours in that we examine incentives to overweight public guidance resulting in higher analyst forecast errors, *ex post*.

$$\times ANALYSTS + \hat{\alpha}_{10}DOWN + \hat{\alpha}_{11}REPUTATION + \hat{\alpha}_{12}AGREE + \hat{\alpha}_{13}HORIZON + \hat{\alpha}_{14}RANGE + \hat{\alpha}_{15}ANALYSTS;$$ and

(3) the residual from the estimation:

$$RESIDUAL\ REVISION = REVISION - SEO\ REVISION - STANDARD\ REVISION.$$

Second, we regress the absolute value of the *ex post* analyst forecast error on the absolute value of the decomposed revision components to test if analyst forecast errors are larger as a result of the increased reliance on management guidance prior to equity offerings:

$$\begin{aligned} |AFE| = {} & \alpha_0 + \alpha_1|SEO\ REVISION| + \alpha_2|STANDARD\ REVISION| + \alpha_3|RESIDUAL\ REVISION| \\ & + \alpha_4|\Delta EPS| + \alpha_5 LOSS + \alpha_6 SPECIAL\ ITEMS + \alpha_7 ANALYSTS + \alpha_8 HORIZON \\ & + \alpha_9 RANGE + \alpha_{10} ANALYST\ DISPERSION + \alpha_{11} FIRM\ SIZE + \alpha_i \sum_{i=1994}^{2005} \text{YEAR}_i + \varepsilon \end{aligned} \quad (2)$$

where $|AFE|$ is the absolute value of the revised consensus analyst forecast error (the difference between actual earnings and the revised analyst consensus forecast), scaled by the stock price at the beginning of the quarter.[16]

We consider eight variables to control for the innate difficulty in estimating earnings, which might vary systematically for firms issuing equity. We also include yearly indicator variables, since analyst forecast errors have changed over time (Brown 2001). To control for unexpected earnings shocks, we include $|\Delta EPS|$, the absolute change in earnings per share from First Call, scaled by the stock price at the beginning of the quarter. Abarbanell and Lehavy (2002) document that forecast errors are larger for firms with losses and special items. Thus, we include *LOSS*, an indicator variable that equals 1 if earnings per share from First Call is negative. and 0 otherwise, and *SPECIAL ITEMS* from Compustat (#32), scaled by the market value of common stock at the beginning of the quarter. We control for *ANALYSTS*, the log of the number of analyst forecasts used to form the revised consensus analyst forecast, as the number of analysts is expected to reduce the *ex post* analyst forecast error (Brown 1997). The next control is *HORIZON*, the log of the number of days between the issuance of management guidance and the earnings announcement. (Recall that the consensus is measured within 15 days of this event; thus, *HORIZON* based on management guidance is highly correlated with the age of the consensus forecast.) The earlier in the quarter the analyst forecast is issued, the greater is the expected error (as there is more time for innovations in earnings to occur). We include two proxies for uncertainty. First, *RANGE* is based on the precision of the management guidance; the greater the *RANGE*, the greater the inferred uncertainty on the part of management, and the larger the expected forecast error. Second, *ANALYST DISPERSION* represents the uncertainty among analysts about forecasted earnings. This variable is the standard deviation of the individual analyst forecasts used to generate the revised consensus analyst forecast, scaled by the stock price at the beginning of the quarter. Finally,

[16] By taking the absolute value of the revision, we focus on the degree of incorporation, the variable of interest, instead of the direction of incorporation. Consider two firms. The manager of Firm A issues guidance that is ten cents below the analyst consensus, and the analysts revise the consensus forecast downward by 0.08. The manager of Firm B issues guidance that is ten cents above the analyst consensus, and the analysts revise the consensus upward by 0.08. *REVISE* is −0.08 and 0.08 for Firm A and Firm B, respectively; however, in both instances the analysts incorporate 80 percent of the guidance.

because larger firms tend to have more disclosure in general, we control for *FIRM SIZE*, the log of market value of common stock at the beginning of the quarter.

A positive coefficient on *SEO REVISION* would indicate that the increased incorporation of guidance by analysts prior to equity offerings results in larger *ex post* forecast errors. We expect the coefficient on *STANDARD REVISION* to be negative, as, in general, management guidance has been shown to lower analyst forecast errors (e.g., Hassell and Jennings 1986). We do not predict the sign of the coefficient on the residual analyst revision.

Table 5 presents the results from estimating Regression (2). Consistent with the increased weighting of management guidance prior to equity offerings representing a tool to please management, it is associated with higher analyst forecast errors, as the coefficient on *SEO REVISION* is positive and significant with and without control variables. In terms of economic magnitude, the overweighting of management guidance prior to equity offerings increases *ex post* forecast errors by 40 percent, on average.[17] *STANDARD REVISION*, the revision associated with known factors, is positive and weakly significant when omitting control variables, but negative and significant when control variables are included.[18] The latter result is consistent with management guidance generally improving analyst forecast accuracy. Finally, *RESIDUAL REVISION* is positively associated with analyst forecast errors across both specifications. This finding suggests that other incentives, in addition to underwriting business, might lead analysts to vary their incorporation of management guidance at the cost of *ex post* accuracy; we examine additional incentives in Section V.

The coefficients on the control variables are largely consistent with expectations. Large shocks to earnings decrease forecast accuracy. Losses and special items are not significant, although special items are generally excluded from core earnings determined by forecast data providers (Philbrick and Ricks 1991). The more analysts following the firm, the greater the forecast accuracy, consistent with Brown (1997). When the management guidance is older or less precise, the *ex post* analyst forecast errors are larger, and the errors are also larger when there is less of a consensus among analysts (i.e., a higher standard deviation of the individual forecasts). Finally, firm size is not significant, which is reasonable given that the sample is restricted to firms that provide guidance, which tend to be larger firms.

Overall, we find that analysts' abnormally high weighting of management guidance preceding equity offerings is associated with an increase in analyst forecast errors. Analysts appear to be sacrificing their forecast accuracy by overweighting management guidance prior to equity offerings.

Cross-Sectional Tests of Analysts' Overweighting of Management Guidance

Thus far, we use consensus forecasts issued by all analysts. Not all analysts, however, are vying for investment bank business. In this section, we examine whether this behavior differs by analyst type. Specifically, we expect dependent analysts who *can* acquire the investment banking business to overweight management guidance when revising their short-term forecasts prior to an equity offering. We do not expect, however, to observe independent analysts overweighting guidance in a similar fashion. We also expect the overweighting of guidance to be associated with the

[17] To determine the economic magnitude of the increased forecast error associated with the overweighting of guidance prior to equity offerings, we multiply the average *SEO REVISION* for SEO firms (0.0004; not tabulated) with the coefficient on *SEO REVISION* (1.2065), which indicates an increase in the analyst forecast error of 0.00048 on average, or 40 percent of the mean analyst forecast error (0.0012; see Table 3).

[18] The coefficient on *SEO REVISION* (1.2065) appears to be much larger than the coefficient on *STANDARD REVISION* (0.0368). In terms of economic significance, however, the importance of *STANDARD REVISION* is about double that of *SEO REVISION*, where we measure economic significance as the change in $|AFE|$ when the *REVISION* variable increases by one standard deviation.

TABLE 5

***Ex Post* Analyst Forecast Accuracy**

		Dependent Variable = |*AFE*|	
Independent Variables	**Predicted Sign**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
Intercept		0.0006***	−0.0038***
		(10.52)	(−4.89)
|*SEO REVISION*|	+	1.3837**	1.2065**
		(2.22)	(2.26)
|*STANDARD REVISION*|	−	0.0198	−0.0368**
		(1.57)	(−2.53)
|*RESIDUAL REVISION*|	?	0.6219***	0.4278***
		(8.86)	(5.60)
|Δ*EPS*|	+		0.0183***
			(4.56)
LOSS	+		0.0001
			(0.96)
SPECIAL ITEMS	+		−0.0003
			(−0.31)
ANALYSTS	−		−0.0002***
			(−3.05)
HORIZON	+		0.0012***
			(6.85)
RANGE	+		0.1018**
			(2.54)
ANALYST DISPERSION	+		0.2489***
			(4.10)
FIRM SIZE	−		−0.0000
			(−1.26)
Year Indicators		Not Included	Included
Number of SEO Observations		87	60
Total Number of Observations		2,997	2,342
Adjusted R^2		0.1500	0.2762

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test.

\|*AFE*\| is the absolute value of the revised analyst forecast error. We decompose the total analyst revision, from the estimation of Regression (1), into three components: that associated with the upcoming equity offering ($SEO\ REVISION = \hat{\alpha}_2 SEO + \hat{\alpha}_9 MGUIDE \times SEO$), that associated with other known factors affecting analyst revisions ($STANDARD\ REVISION = \hat{\alpha}_0 + \hat{\alpha}_1 MGUIDE + \hat{\alpha}_3 MGUIDE \times DOWN + \hat{\alpha}_4 MGUIDE \times REPUTATION + \hat{\alpha}_5 MGUIDE \times AGREE + \hat{\alpha}_6 MGUIDE \times HORIZON + \hat{\alpha}_7 MGUIDE \times RANGE + \hat{\alpha}_8 MGUIDE \times ANALYSTS + \hat{\alpha}_{10} DOWN + \hat{\alpha}_{11} REPUTATION + \hat{\alpha}_{12} AGREE + \hat{\alpha}_{13} HORIZON + \hat{\alpha}_{14} RANGE + \hat{\alpha}_{15} ANALYSTS$), and the residual from the estimation ($RESIDUAL\ REVISION = REVISE - SEO\ REVISION - STANDARD\ REVISION$). All t-statistics are based on robust standard errors clustered at the firm level.

Additional variables are defined in Table 2.

ex post selection of the underwriter and, thus, we expect those analysts whose employers are *ex post* selected to underwrite the deal to have incorporated the greatest proportion of guidance into their short-term earnings forecasts, after controlling for the credibility and usefulness of the guidance. For these analyses, we limit the sample to guidance issued in the year prior to the announce-

ment of an equity offering (SEO observations). As a result, these cross-sectional tests also help to allay concerns that some underlying difference between SEO and non-SEO observations drives the higher weighting of guidance among SEO observations.

For each analyst following the 202 firms that issue management guidance and announce an equity offering in the subsequent year, we identify (1) whether they are employed by an investment bank, and (2) whether their employer is on the subsequently formed team of underwriters; in these instances, we set *DEPENDENT* and *AFFILIATED*, respectively, equal to 1.[19] Approximately 90 (25) percent of the analysts in the sample are dependent (affiliated). To be included in the sample, the firm must be followed by both dependent and independent (affiliated and unaffiliated) analysts. This design choice rules out the possibility that different types of analysts tend to follow different types of firms (and therefore different-quality management guidance). This requirement results in 35 (172) usable management guidance observations for the dependent (affiliated) analyst-type tests.[20]

For each instance of management guidance, we have two observations, one for each analyst group (i.e., dependent and independent), resulting in a total of 70 (344) observations for the dependent (affiliated) test. For each of these subgroups, we generate *MGUIDE* and *REVISE* using the corresponding median consensus forecast that includes only the relevant analysts (dependent/independent and affiliated/unaffiliated). The reason for this approach is that when we examine affiliated/unaffiliated or dependent/independent analysts, each of these four sub-groups will have a potentially different pre-existing and revised consensus, as each analyst group's consensus should be made up of only that particular analyst type (e.g., there will be an affiliated analyst consensus, which might differ from the unaffiliated analyst consensus). Partitioning in this manner allows us to determine whether the main movement in the consensus (documented in Table 4) is by self-interested (dependent or affiliated) analysts, supporting our hypothesis, or by all analysts, supporting the existence of a correlated omitted variable or self-selection concerns.

We replicate the estimation of Regression (1), but instead of investigating variation based on *SEO*, we hypothesize that there will be variation based on analyst type (recall that the test is limited to the SEO firms). Table 6 presents the results from this estimation, with dependent/independent analysts in the first column and affiliated/unaffiliated in the second column.

Turning first to the dependent/independent partition, we find that, conditional on an upcoming announcement of an equity offering, those analysts who can gain the underwriting business incorporate significantly more of the management guidance into their short-term earnings forecasts than do independent research analysts. Specifically, the coefficient on *MGUIDE* × *DEPENDENT* is positive and significant, while the coefficient on *MGUIDE* is not significantly different from zero. Unlike dependent analysts, who have an incentive to gain the subsequent underwriting business, independent analysts do not appear to incorporate any management guidance beyond that dictated by the credibility and usefulness of the guidance.[21] Thus, the overweighting of management guidance preceding equity offerings that we document in Table 4 is, on average, present only among analysts with greater incentives to please management—those who can actually obtain the

[19] To determine if the analysts' employers could feasibly gain the underwriting business, we use the data from Gu and Xue (2008), who separate full service investment banks and brokerage firms from independent research firms; we thank Zhaoyang Gu for providing us with these data. Ideally, we would like to focus solely on those analysts employed by banks that offer underwriting services (as some brokerage firms do not). We are limited, however, to this broader partition, which biases against finding results consistent with our hypothesis. We use SDC Platinum to identify the subsequent underwriting team. Analysts are affiliated if they are employed by any investment bank involved in the deal, rather than solely by the lead underwriter.

[20] To maximize the number of usable observations, we set *REPUTATION* equal to 0 if it is missing.

[21] If we exclude the control variables, the coefficient on the main effect of *MGUIDE* is positive and significant, consistent with independent research analysts incorporating management guidance into their earnings forecasts, on average.

TABLE 6

Revision Magnitude by Analyst Type

		Dependent Variable = *REVISE*	
		Analyst Type	
		Dependent Analysts	Affiliated Analysts
Independent Variables	**Predicted Sign**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
Intercept		−0.0006	−0.0029*
		(−0.44)	(−1.85)
MGUIDE	+	1.4482	0.7540
		(1.32)	(0.61)
MGUIDE × *ANALYST TYPE*	+	0.2387*	0.3355*
		(1.70)	(1.74)
MGUIDE × *DOWN*	+	−0.3762*	−0.2403
		(−1.88)	(−1.28)
MGUIDE × REPUTATION	+	−0.0846	−0.0446
		(−0.81)	(−0.39)
MGUIDE × *AGREE*	+	0.5982***	−0.0235
		(2.90)	(−0.11)
MGUIDE × *HORIZON*	+	−0.2661	−0.1785
		(−1.23)	(−0.58)
MGUIDE × *RANGE*	−	24.2946	−5.5979
		(1.10)	(−0.32)
MGUIDE × *ANALYSTS*	?	−0.1454	0.4167***
		(−1.17)	(3.60)
Main Effects		Included	Included
Analyst Type Equal to 1		35	172
Total Number of Observations		70	344
Adjusted R^2		0.7945	0.7487

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test.

Only firms issuing equity in the next year are included in these estimations. We define *DEPENDENT* (*AFFILIATED*) as an indicator variable that is set equal to 1 if the analyst is employed by a full-service investment bank or a brokerage house (the subsequent underwriter), and 0 otherwise. We recalculate *MGUIDE* and *REVISE*, defined in Table 2, for both the dependent/affiliated and independent/unaffiliated analysts' initial forecasts and subsequent revisions (based on the median consensus analyst forecast for that specific type of analyst; we examine only firm-quarters that have both types of analyst forecasts). All t-statistics are based on robust standard errors clustered at the firm level.

Additional variables are defined in Table 2.

underwriting business for their investment bank. This finding is consistent with analysts trying to please management by overweighting their guidance.

Turning next to the affiliated/unaffiliated partition, we again find that the coefficient on *MGUIDE* is insignificant, while the coefficient on *MGUIDE* × *AFFILIATED* is positive and significant. These results suggest that the analysts who incorporate the greatest amount of the management guidance are affiliated analysts—those who work for the investment banks housing the subsequent underwriters of the equity offering. These results are consistent with analysts benefiting from overweighting management guidance. Specifically, although the overweighting of

guidance is associated with lower analyst forecast accuracy (Table 5), it is also associated with future underwriting business.[22]

Overall, the results are consistent with our central hypothesis: analysts appear to overweight management guidance when revising their short-term earnings forecasts in order to please management. This behavior precedes the announcements of equity offerings, leads to larger *ex post* analyst forecast errors, is present only among those analysts who can gain the investment banking business, and is associated with the *ex post* receipt of the underwriting services. The following section presents additional analyses to corroborate our conclusions.

V. ADDITIONAL ANALYSES

Self-Selection of an Equity Offering

Although we attempt to control for the underlying economics of each firm in each of the tests, and the cross-sectional tests help alleviate the endogeneity concerns, in this section we explore the issue of endogeneity further. First, we employ the "omitted variables" variant of the Hausman (1978, 1983) test to determine if the main treatment variable, *MGUIDE* × *SEO*, exhibits evidence of endogeneity after we control for the variables presented in Table 4. The test is unable to reject the null condition of no endogeneity ($p = 0.31$; not tabulated).

Second, we control for self-selection bias using a two-stage approach in which the first stage is a probit regression of *SEO* on the hypothesized determinants of equity offerings (presented in the Appendix). This first-stage regression uses the following expected determinants (with annual data): profitability, R&D and capital expenditure investments, the book-to-market ratio, leverage, firm size, cash reserves, tangible assets, annual stock returns, financial stress, and equity offering history (Rajan and Zingales 1995; Jung et al. 1996; Loughran and Ritter 1997; Frank and Goyal 2003; Dittmar and Thakor 2007; Kadan et al. 2009). From this first-stage regression, which identifies the likelihood of a firm choosing to issue equity, we calculate the inverse Mills ratio (see Heckman 1979; Leuz and Verrecchia 2000) and include it in Regression (1). Including this ratio helps control for the likelihood of self-selecting into the equity issuer group. The results are similar to those in Table 4 when we include the inverse Mills ratio (not tabulated).

Finally, following Frank (2000) and Larcker and Rusticus (2010), we describe the characteristics an omitted variable must have in order to overturn the OLS results. The unidentified omitted variable must be correlated with both *REVISE* (the dependent variable) and *MGUIDE* × *SEO* (the treatment variable) in order to cause spurious positive correlations between *REVISE* and *MGUIDE* × *SEO*. Specifically, the product of two partial correlations, the correlation between the omitted variable and *REVISE*, and the correlation between the omitted variable and *MGUIDE* × *SEO*, needs to be at least 0.0415 to overturn the positive association between *REVISE* and *MGUIDE* × *SEO*. To put this number in perspective, the highest product of the partial correlations with *REVISE* and *MGUIDE* × *SEO* based on the identified control variables is 0.0299 (for *MGUIDE* × *REPUTATION*). Thus, we would need an omitted variable with an impact 39 percent greater than that of any of the control variables to change the results, suggesting that it is unlikely that a confounding omitted variable would overturn the positive association between *REVISE* and *MGUIDE* × *SEO*.

[22] It is possible that affiliated analysts are simply more diligent at following the firm and incorporating new information about the firm (along the lines of Jacob et al. 2008). This diligent incorporation, however, should not be associated with larger *ex post* analyst forecast errors. Alternatively, some analysts might always wish to please management. We also, therefore, examine the incorporation of guidance by affiliated versus unaffiliated analysts in all years that do not precede an equity offering. Unlike in the year prior to the equity offering, we do not find a difference in the incorporation of guidance between these two types of analysts when guidance is issued at other times (not tabulated), suggesting that analysts appear to change their behavior when equity offerings are imminent.

To summarize, we are unable to reject the null of no endogeneity based on the omitted variables variant of the Hausman test, results are similar after including the inverse Mills ratio, and the likelihood that our OLS results are driven by an omitted variable is low. Regardless, we cannot completely rule out the possibility that the endogenous nature of equity offerings drives our results.

Alternative Mechanisms to Please Management

We document a tool that analysts appear to use to please management: the overweighting of management guidance by analysts revising their short-term earnings forecasts. Prior research has found evidence that analysts issue overly optimistic stock recommendations (e.g., Dugar and Nathan 1995; Lin and McNichols 1998; Michaely and Womack 1999) for firms they wish to please.[23] It is likely that analysts use a portfolio of tools to curry favor with management. To investigate the interplay of these tools, we identify the analysts in the sample who issue a stock recommendation, and partition those issuing a strong buy recommendation from those issuing any other recommendation (buy, hold, or sell). To be included in the test, each guidance observation must have at least one strong buy recommendation and one other recommendation. As with the analyst-type tests, we re-estimate *REVISE* and *MGUIDE* for both groups of analysts. We estimate Regression (1) separately for the two groups (strong buy and other), presented in Table 7. We find that only analysts who issue a strong buy recommendation overweight management guidance prior to an equity offering ($\alpha_2 = 0.18$, $p = 0.08$). In contrast, among analysts not issuing a strong buy recommendation, there is no evidence of analysts overweighting guidance prior to an equity offering for the same management guidance observations ($\alpha_2 = -0.04$, $p > 0.50$). These findings suggest that analysts please managers by jointly using both optimistic stock recommendations and the overweighting of guidance.[24] These findings also further establish that analysts with different incentives respond differently to the same guidance, which is more consistent with the overweighting of guidance representing a mechanism used to curry favor with management and less consistent with alternative explanations such as the endogeneity of equity offerings.

Knowledge of the Equity Offering

An assumption underlying our study is that analysts know which firms will be issuing equity in the upcoming year and thus actively vie for the underwriting business. We conduct two tests to determine the robustness of the results to this assumption. First, we replace actual equity announcements from SDC with an expected probability that the firm will issue equity in the year to come (see the Appendix). We find that analysts incorporate more guidance for firms that are *expected* to have an equity offering, suggesting that analysts make overtures to all management teams that *could* be undertaking an equity offering in the near future (not tabulated).

Second, we examine the sensitivity of the results to the length of the horizon period. Specifically, we examine announcements that occur in the six-month period following the management guidance relative to those that occur in the 7–12-month period following the management guidance. If the incentive to gain underwriting business is driving the results, then we expect stronger results when the horizon is shorter—i.e., when equity offerings are both more certain and have more timely payoffs. Alternatively, if firms that tend to issue equity also tend to provide more useful guidance, we would not expect to find stronger results in the shorter time horizon.

---

[23] Research has also documented overly optimistic long-term growth forecasts as a mechanism to please management (e.g., Lin and McNichols 1998; Dechow et al. 2000); however, this variable is more difficult to identify, *ex ante*. Thus, we focus the cross-sectional test on stock recommendations.

[24] Results are similar if we partition only those firms announcing equity offerings in the subsequent year by stock recommendation. *MGUIDE* is positive and significant only among the strong buy partition (not tabulated).

TABLE 7

Revision Magnitude on Equity Offerings Based on Stock Recommendation and Horizon

		Dependent Variable = *REVISE*			
		Analyst Recommendation		Management Guidance Issued	
		Buy, Hold, or Sell	Strong Buy	In the Six Months before the SEO Announcement	In the 7–12 Months before the SEO Announcement
Independent Variables	**Predicted Sign**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
Intercept		0.0001	0.0011	−0.0012*	−0.0014**
		(0.16)	(1.05)	(−1.75)	(−1.97)
MGUIDE	+	−0.3200	−0.2230	−0.1627	−0.2287
		(−0.68)	(−0.53)	(−0.37)	(−0.50)
MGUIDE × SEO	+	−0.0365	0.1790*	0.3287***	0.0431
		(−0.29)	(1.75)	(4.29)	(0.72)
MGUIDE × DOWN	+	0.1082	0.0512	0.1413*	0.1430*
		(1.27)	(0.72)	(1.88)	(1.87)
MGUIDE × REPUTATION	+	0.1239***	0.0835*	0.1498***	0.1494***
		(2.83)	(1.80)	(3.04)	(3.08)
MGUIDE × AGREE	+	0.1866***	−0.0169	0.1571***	0.1711***
		(2.73)	(−0.27)	(2.98)	(3.15)
MGUIDE × HORIZON	+	0.0365	0.0062	0.1077	0.1183
		(0.27)	(0.06)	(1.05)	(1.12)
MGUIDE × RANGE	−	16.5235**	19.5753*	−8.3088***	−7.4736**
		(2.21)	(1.80)	(−2.61)	(−2.38)
MGUIDE × ANALYSTS	?	0.2210***	0.1485***	0.1155***	0.1172***
		(4.22)	(3.01)	(2.88)	(2.92)
Main Effects		Included	Included	Included	Included
Number of SEO Observations		59	59	45	42
Total Number of Observations		2,133	2,133	2,983	2,978
Adjusted R^2		0.7374	0.3548	0.8806	0.8733

*, **, *** Indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test.
All t-statistics are based on robust standard errors clustered at the firm level.
Additional variables are defined in Table 2

Table 7 presents the results in the final two columns, with the shorter time horizon announcements in the first column and the 7–12-month time horizon announcements in the second column. Consistent with our expectations, analysts weight management guidance more heavily when the equity offerings are nearer term in nature ($\alpha_2 = 0.33$; $p < 0.01$). When the equity offering is not imminent, there is no evidence, on average, of analysts overweighting management guidance ($\alpha_2 = 0.04$; $p = 0.47$), consistent with analysts having less certainty about the possibility of an offering or placing a lower weight on the less timely benefits.

Upward versus Downward Guidance

Chan et al. (2003b), Ke and Yu (2006), Baik and Yi (2007), and Malmendier and Shanthikumar (2007) generally find that affiliated analysts tend to issue more pessimistic short-term forecasts, thereby allowing managers to meet the analyst forecast, and Richardson et al. (2004) document a "walk down" of analyst forecasts that is more pervasive prior to equity offerings. It is likely that the overweighting of guidance we document is one way that managers walk down analysts to *ex post* pessimistic analyst forecasts (e.g., Cotter et al. 2006). However, we posit that managers will also, at times, wish analysts to increase their forecasts.[25] Generally, we expect analysts to take their cue from management, regardless of the direction of the management guidance.

We investigate this directly by examining the direction of the guidance in Table 8. Of the 87 instances of management guidance issued in the year before the announcement of the equity offering with all necessary control variables, 31 are upward, 42 are downward, and 14 are confirmations. We re-estimate Regression (1) separately for upward and downward guidance and find that analysts overweight management guidance prior to equity offerings in both groups. Moreover, the magnitude of the overweighting appears similar for upward and downward forecasts (0.31 versus 0.33), suggesting that analysts are willing to overweight management guidance by an economically significant amount in either direction to please management. Thus, our findings go beyond those documenting a "walk down" of analysts, as we find analysts are also "walked up."

Additional Settings

Though we examine the specific setting of equity offerings, the overweighting of management guidance is also a viable tool for analysts wishing to curry favor with management in the face of other incentives. As a final robustness check, we examine two additional incentives, the desire to gain debt underwriting fees and M&A advisory fees. These incentives are present in the year following management guidance for 6.3 and 4.0 percent of the sample firms, respectively. We examine all three incentives (*SEO*, *DEBT*, and *M&A*) in the first column of Table 9, and we create a joint indicator variable (*INCENTIVES*) that equals 1 if any of these three incentives are present in the second column of Table 9. Turning first to the disaggregated estimation, we see that the coefficient on *MGUIDE* × *SEO* is 0.23 ($p < 0.01$), similar to the results in Table 4. Analysts also overweight management guidance prior to debt offerings; the coefficient on *MGUIDE* × *DEBT* is 0.13 ($p = 0.02$). Finally, there is some evidence that analysts overweight guidance when there are prospective M&A advisory services; the coefficient on *MGUIDE* × *M&A* is 0.11 ($p = 0.12$). Thus, the results are not limited to the incentives of equity offerings, but rather extend to additional settings where analysts might gain fees for their banks. When we aggregate all of the incentives, 342 firm-year observations (11.3 percent of the sample) have at least one of these three incentives in the following year. The coefficient on *MGUIDE* × *INCENTIVES* is 0.16 ($p = 0.02$), consistent with the premise that analysts with incentives to please management overweight management guidance when revising their short-term earnings forecasts.

In summary, we consider the endogeneity of equity offerings, the interaction of different mechanisms to curry management favor (e.g., stock recommendations), the knowledge of the offering, the direction of the management guidance, and alternative measures of analysts' incentives (e.g., debt offerings).[26] Each of these results supports our conclusion that analysts wishing to please management overweight management guidance when revising their short-term earnings forecasts.

---

[25] Evidence of downward management guidance is not direct evidence of *ex post* pessimism; however, finding that managers provide upward guidance suggests that they do not desire pessimistic forecasts alone (or they would simply allow the analysts to underestimate earnings).

[26] Results are also similar if we limit the sample to firms that announce an equity offering at least once during the sample period (not tabulated). We also split the sample on whether the management guidance occurred before or after Regu-

TABLE 8

Revision Magnitude on Upward and Downward Management Guidance

		Dependent Variable = *REVISE*	
		Coefficient (t-statistic)	Coefficient (t-statistic)
Independent Variables	**Predicted Sign**	**Upward Management Guidance**	**Downward Management Guidance**
Intercept		−0.0001	−0.0037
		(−0.08)	(−1.45)
MGUIDE	+	0.0020	−0.0853
		(0.00)	(−0.14)
MGUIDE × SEO	+	0.3059**	0.3269***
		(2.14)	(3.50)
MGUIDE × REPUTATION	+	0.0621	0.1792***
		(0.45)	(2.93)
MGUIDE × AGREE	+	0.3882**	0.1138*
		(2.37)	(1.67)
MGUIDE × HORIZON	+	0.0769	0.1134
		(0.32)	(0.84)
MGUIDE × RANGE	−	−24.8156**	−6.4814*
		(−2.29)	(−1.76)
MGUIDE × ANALYSTS	?	0.0487	0.1397***
		(0.51)	(2.59)
Main Effects		Included	Included
Number of SEO Observations		31	42
Total Number of Observations		865	1,514
Adjusted R^2		0.6202	0.8503

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test.

Management guidance is upward if the guidance is greater than the preexisting consensus analyst forecast and downward if the guidance is less than the preexisting consensus analyst forecast. All t-statistics are based on robust standard errors clustered at the firm level.

Additional variables are defined in Table 2.

VI. CONCLUSION

We predict and find that analysts who wish to please the firms they follow overweight management earnings guidance when revising their short-term earnings forecasts, even at the expense of *ex post* accuracy. We also document evidence of the benefit to analysts—subsequent underwriting business for their firms. Specifically, we show that analysts who are affiliated with the *ex post* underwriters of the subsequent equity offerings weight guidance more heavily than other analysts.

lation Fair Disclosure in October 2000, and the Global Research Analyst Settlement in April 2003. In results not tabulated, we find a significantly stronger association between management guidance and analysts' incorporation of that guidance preceding equity offerings in the post-Regulation Fair Disclosure era, when other mechanisms of guidance were restricted, and we find an economically weaker relation between upcoming equity offerings and analysts' revisions following the settlement, a period when analysts faced increased scrutiny by regulators.

TABLE 9

Analyst Revision Magnitude on Multiple Incentives: Equity Offerings, Debt Offerings, and M&A Activity

		Dependent Variable = *REVISE*	
Independent Variables	**Predicted Sign**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
Intercept		−0.0012*	−0.0011*
		(−1.78)	(−1.68)
MGUIDE		−0.1722	−0.0746
		(−0.39)	(−0.14)
MGUIDE × SEO	+	0.2293***	
		(3.96)	
MGUIDE × DEBT	+	0.1327**	
		(2.30)	
MGUIDE × M&A	+	0.1062	
		(1.57)	
MGUIDE × INCENTIVES	+		0.1609**
			(2.40)
MGUIDE × DOWN	+	0.1439*	0.1516**
		(1.93)	(2.04)
MGUIDE × REPUTATION	+	0.1418***	0.1380***
		(2.89)	(2.71)
MGUIDE × AGREE	+	0.1685***	0.1576***
		(3.12)	(2.82)
MGUIDE × HORIZON	+	0.1076	0.0871
		(1.05)	(0.86)
MGUIDE × RANGE	−	−7.3614**	−6.6200**
		(−2.48)	(−2.22)
MGUIDE × ANALYSTS	?	0.1155***	0.1130***
		(2.88)	(2.80)
Main Effects		Included	Included
Number of SEO Observations		87	87
Number of DEBT Observations		175	175
Number of M&A Observations		115	115
Number of Incentives Observations		342	342
Total Number of Observations		3,025	3,025
Adjusted R^2		0.8810	0.8789

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test.

INCENTIVES is an indicator variable that is set equal to 1 if there is an equity offering, debt offering, or M&A activity in the year following the management guidance, and 0 if none of these events occur in the subsequent year. Standard errors are clustered by firm. All t-statistics are based on robust standard errors clustered at the firm level.

Additional variables are defined in Table 2.

Our study has limitations. First, equity offerings are endogenous, and characteristics of firms issuing equity might lead analysts to incorporate more guidance prior to equity offerings for reasons other than gaining underwriting business. To minimize this limitation, we explicitly link the overweighting of guidance prior to equity offerings to larger *ex post* forecast errors, mitigating the possibility that the overweighting is in response to more informative guidance prior to offer-

ings. We also examine different analysts' responses to the same guidance and find that only analysts with an incentive to please management weight the management guidance more than predicted based on the credibility and usefulness of the guidance. In addition, we conduct a Hausman test that does not reject the null of no endogeneity. Finally, we estimate a Heckman two-stage regression to further control for the endogenous choice to provide equity and continue to find similar results. A second limitation is that we do not model the manager's selection of an underwriting team. Clearly, managers will consider deal terms, underwriter quality, and other criteria when determining whom to select as underwriters. This analysis is beyond our scope. These additional criteria, however, would likely add noise to the analysis. A third limitation is that all of the main tests and inferences surround a single incentive: upcoming equity offerings. We therefore conduct a preliminary analysis on both debt issues and M&A activity, where analysts might wish to vie for the debt underwriting business or the M&A advisory fees. We find evidence of analysts overweighting guidance prior to these events as well.

Overall, our findings suggest that analysts weight management guidance more heavily than expected when faced with incentives to please management. Future research might examine other settings where analysts might overweight management guidance, such as pre-Regulation Fair Disclosure selective disclosure (Chen and Matsumoto 2006; Ke and Yu 2006) or favored conference call participation (Mayew 2008). Future research might also investigate whether investors anticipate that analysts overweight guidance in certain settings, whether analysts' reputations are hurt by this behavior, or whether analysts appear to understand individual managers' incentives and incorporate guidance differentially based on these perceived incentives.

APPENDIX

We form an SEO prediction model to test two of our robustness checks: the choice to issue equity and the knowledge of the upcoming equity offering versus the expectation of such an event (both presented in Section V). Our model is as follows:

$$\begin{aligned} SEO_{t+1} = {} & \alpha_0 + \alpha_1 PROFITABILITY_t + \alpha_2 \Delta SALES_t + \alpha_3 INVESTMENT_t + \alpha_4 B/M\ RATIO_t \\ & + \alpha_5 LEVERAGE_t + \alpha_6 SIZE_t + \alpha_7 CASH_t + \alpha_8 TANGIBLE\ ASSETS_t \\ & + \alpha_9 ANNUAL\ RETURN_t + \alpha_{10} FINANCIAL\ STRESS_t + \alpha_{11} SEO\ HISTORY \\ & + \text{Industry Indicators} + \text{Year Indicators} + \varepsilon_t. \end{aligned} \tag{A1}$$

Each of the variables is defined in Table 10. We are predicting actual equity offerings in year $t+1$ and include the following expected determinants from year t, based on prior research (e.g., Rajan and Zingales 1995; Jung et al. 1996; Loughran and Ritter 1997; Frank and Goyal 2003; Dittmar and Thakor 2007; Kadan et al. 2009).

First, we include firm profitability. Generally, the more profitable the firm, the more likely it is to issue equity, relative to debt. However, we are predicting an equity offering relative to no equity offering (instead of the use of debt). Thus, it could be that more profitable firms have a lower requirement for additional funds and, thus, we have no prediction on the sign of profitability. Our next variable, change in sales ($\Delta SALES$), is a proxy for growth, and we expect growth firms to be more likely to issue equity. *INVESTMENT* is the amount of expenditures on R&D and capital investments in year t; generally, the higher these expenditures, the higher the expected probability of an equity offering. We expect the book-to-market ratio (*B/M RATIO*) to be negatively related to equity offerings; again, higher growth firms are expected to have a higher probability of issuing equity. Leverage is expected to increase a firm's likelihood of issuing equity over debt, but as noted above, we are predicting the likelihood of issuing equity versus not issuing equity and, thus, we have no prediction for the associated variable. We include *SIZE* (the log of sales), as larger

TABLE 10
SEO Prediction Model

Independent Variables	Predicted Sign	Dependent Variable = *SEO* Coefficient (χ^2-statistic)
Intercept		−1.5554*** (22.45)
PROFITABILITY	?	−0.2069** (4.21)
ΔSALES	+	0.2309*** (38.67)
INVESTMENTS	+	0.4918*** (11.90)
B/M RATIO	−	−0.3305*** (37.20)
LEVERAGE	?	0.9946*** (81.33)
SIZE	−	−0.0696*** (39.18)
CASH	−	−0.1662*** (13.76)
TANGIBLE ASSETS	−	0.1453** (3.85)
ANNUAL RETURN	+	0.2831*** (170.08)
FINANCIAL STRESS	+	0.1490** (5.06)
SEO HISTORY	+	0.2878*** (48.76)
Industry Indicators		Included
Year Indicators		Included
Number of SEO Observations		1,104
Total Number of Observations		14,407
Pseudo R^2		0.1435

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under a two-tailed test.

Variable Definitions:

SEO_{t+1} = indicator variable that is set equal to 1 if the firm issued equity in year t+1, and 0 otherwise;
$PROFITABILITY_t$ = profitability ((Annual Compustat #13 + #62)/lagged #6);
$\Delta SALES_t$ = change in sales ((#12 – lagged #12)/lagged #12);
$INVESTMENT_t$ = investments in research and development and capital expenditures in year t (#128 + #46)/lagged #6;
$B/M\ RATIO_t$ = book-to-market ratio (#60/(#54 × #199));
$LEVERAGE_t$ = debt leverage (#9/#6);
$SIZE_t$ = firm size, measured using the log of sales (#12);
$CASH_t$ = cash reserves (#1/#6);
$TANGIBLE\ ASSETS_t$ = measure of collateral or tangible assets (#8/lagged #6);
$ANNUAL\ RETURN_t$ = market-adjusted fiscal year return for year t;
FINANCIAL STRESS = sum of cash dividends, net investments, and change in working capital less cash inflows from operations in year t (see Frank and Goyal 2003); and
SEO HISTORY = 1 if a firm issued equity in previous three years, and 0 otherwise.

firms are expected to have lower costs of debt and thus may be more likely to issue debt. Firms with more cash are expected to be less likely to issue equity, while firms with greater tangible assets are expected to be more likely to issue debt than equity. We expect firms with higher returns over the past year to be more likely to issue equity; *ANNUAL RETURN* is the market-adjusted fiscal year return for year *t*. We include a variable for financial stress, following Frank and Goyal (2003), who argue that demand for capital is closely tied to the firm's financing deficit. *FINANCIAL STRESS* is a measure of excess cash. Finally, we include *SEO HISTORY*, which is an indicator variable showing whether the firm issued equity in the prior three years.[27]

Table 10 reports results. Generally, the predicted signs are consistent with existing theory, although, as noted above, our setting differs, in that we are predicting the difference between equity offering firms and all other firms; we are not conditioning on the need for additional financing. Overall, the estimated pseudo-R^2 is approximately 14 percent.

We form the inverse Mills ratio following Heckman (1979). To form the predicted probability of an upcoming equity offering, we estimate Regression (A1) in a pooled regression (as presented in Table 10) and retain the predicated value for each firm-year. We then merge the annual prediction of an equity offering with guidance quarters in that year to avoid look-ahead bias, as analysts would need year *t* values to predict year *t*+1 equity offerings and, thus, would not be able to form their predictions until year *t*+1.

REFERENCES

Abarbanell, J., and R. Lehavy. 2002. Differences in commercial database reported earnings: Implications for empirical research. Working paper, The University of North Carolina at Chapel Hill, NC.

Ajinkya, B., and M. Gift. 1984. Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research* 22 (2): 425–444.

Baginski, S., E. Conrad, and J. Hassell. 1993. The effects of management forecast precision on equity pricing and on the assessment of earnings uncertainty. *The Accounting Review* 68: 913–927.

———, and J. Hassell. 1990. The market interpretation of management earnings forecasts as a predictor of subsequent financial analyst forecast revision. *The Accounting Review* 65 (1): 175–190.

———, ———, and M. Kimbrough. 2008. Macro information environment change and the quality of management earnings forecasts. *Review of Quantitative Finance and Accounting* 31 (3): 311–330.

Baik, B., and H. Yi. 2007. Are affiliated analysts more likely than unaffiliated analysts to provide EPS forecasts that management can meet or beat? Working paper, Seoul National University, Seoul, Korea.

Bhojraj, S., R. Libby, and H. Yang. 2008. Practice makes perfect: Do firms and markets learn from past earnings guidance behavior? Working paper, Cornell University, Ithaca, NY.

Bhushan, R. 1989. Firm characteristics and analyst following. *Journal of Accounting and Economics* 11: 255–274.

Bradshaw, M., S. Richardson, and R. Sloan. 2006. The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics* 42 (1–2): 53–85.

Brown, L. 1997. Analyst forecasting errors: Additional evidence. *Financial Analysts Journal* 53 (6): 81–88.

——— 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39: 221–241.

Chan, L., J. Karceski, and J. Lakonishok. 2003a. The level and persistence of growth rates. *The Journal of Finance* 58: 643–684.

———, ———, and J. Lakonishok. 2003b. Analysts' conflict of interest and biases in earnings forecasts. NBER Working Paper No. W9544. Cambridge, MA: NBER.

[27] As noted by Kadan et al. (2009), while about 4 percent of listed firms tend to issue equity in a given year, about 13 percent of firms that issued equity in the prior two years have again issued equity. We also control for industry and year, as these will also likely be associated with a firm's decision to issue equity.

Chen, S., and D. Matsumoto. 2006. Favorable versus unfavorable recommendations: The impact on analyst access to management-provided information. *Journal of Accounting Research* 44 (4): 657–689.

Cotter, J., I. Tuna, and P. Wysocki. 2006. Expectations management and beatable targets: How do analysts react to explicit earnings guidance? *Contemporary Accounting Research* 23 (3): 593–624.

Cowen, A., B. Groysberg, and P. Healy. 2006. Which types of analyst firms are more optimistic? *Journal of Accounting and Economics* 41: 119–146.

Dechow, P., A. Hutton, and R. Sloan. 2000. The relation between analysts' long-term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research* 17 (1): 1–32.

Dittmar, A., and A. Thakor. 2007. Why do firms issue equity? *The Journal of Finance* LXII: 1–54.

Dugar, A., and S. Nathan. 1995. The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations. *Contemporary Accounting Research* 12: 131–160.

Frank, K. 2000. Impact of a confounding variable on a regression coefficient. *Sociological Methods & Research* 29 (2): 147–194.

Frank, M., and V. Goyal. 2003. Testing the pecking order theory of capital structure. *Journal of Financial Economics* 67: 217–248.

Givoly, D., and J. Lakonishok. 1979. The information content of financial analysts' forecasts of earnings. *Journal of Accounting and Economics* 1 (3): 165–185.

Gu, Z., and J. Xue. 2008. The superiority and disciplining role of independent analysts. *Journal of Accounting and Economics* 45: 289–316.

Hassell, J., and R. Jennings. 1986. Relative forecast accuracy and the timing of earnings forecast announcements. *The Accounting Review* 61 (1): 58–75.

———, ———, and D. Lasser. 1988. Management earnings forecasts: Their usefulness as a source of firm-specific information to security analysts. *Journal of Financial Research* XI (4): 303–319.

Hausman, A. 1978. Specification tests in econometrics. *Econometrica* 46: 1251–1271.

——— 1983. Specification and estimation of simultaneous equations models. *Handbook of Econometrics* I: 391–448.

Hayes, R. 1998. The impact of trading commission incentives on analysts' stock coverage decisions and earnings forecasts. *Journal of Accounting Research* 36 (2): 299–320.

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–161.

Hirst, E., L. Koonce, and S. Venkataraman. 2008. Management earnings forecasts: A review and framework. *Accounting Horizons* 22: 315–338.

Hong, H., and J. Kubik. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance* LVIII (1): 313–351.

Hutton, A., and P. Stocken. 2009. Prior forecasting accuracy and investor reaction to management earnings forecasts. Working paper, Boston College, Boston, MA.

Jacob, J., S. Rock, and D. Weber. 2008. Do non-investment bank analysts make better earnings forecasts? *Journal of Accounting, Auditing & Finance* 23 (1): 23–61.

Jung, K., Y. Kim, and R. Stulz. 1996. Timing, investment opportunities, managerial discretion, and the security issue decision. *Journal of Financial Economics* 42: 159–185.

Kadan, O., L. Madureira, R. Wang, and T. Zach. 2009. Conflicts of interest and stock recommendations: The effects of the Global Settlement and related regulations. *Review of Financial Studies* 22 (10): 4189–4217.

Ke, B., and Y. Yu. 2006. The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research* 44: 965–999.

King, R., G. Pownall, and G. Waymire. 1990. Expectations adjustment via timely earnings forecast disclosure: Review, synthesis, and suggestions for future research. *Journal of Accounting Literature* 9: 113–144.

Krigman, L., W. Shaw, and K. Womack. 2001. Why do firms switch underwriters? *Journal of Financial Economics* 60: 245–284.

Larcker, D., and T. Rusticus. 2010. On the use of instrumental variables in accounting research. *Journal of Accounting and Economics* 49 (3): 186–205.

Leuz, C., and R. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38: 91–124.

Libby, R., J. Hunton, H. Tan, and N. Seybert. 2008. Relationship incentives and the optimistic/pessimistic pattern in analysts' forecasts. *Journal of Accounting Research* 46: 173–198.

Lim, T. 2001. Rationality and analysts' forecast bias. *The Journal of Finance* LVI (1): 369–385.

Lin, H., and M. McNichols. 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25: 101–128.

Loughran, T., and J. Ritter. 1997. The operating performance of firms conducting seasoned equity offerings. *The Journal of Finance* 52: 1823–1850.

Lys, T., and S. Sohn. 1990. The association between revisions of financial analysts' earnings forecasts and security price changes. *Journal of Accounting and Economics* 13 (4): 341–363.

Malmendier, U., and D. Shanthikumar. 2007. Do security analysts speak in two tongues? NBER Working Paper No. W13124. Cambridge, MA: NBER.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77: 483–514.

Mayew, W. 2008. Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research* 46: 627–659.

McNichols, M., and P. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* 35: 167–199.

Michaely, R., and K. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12: 653–686.

Mikhail, M., B. Walther, and R. Willis. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* 74: 185–200.

Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.

Philbrick, D., and W. Ricks. 1991. Using Value Line and I/B/ E/S analyst forecasts in accounting research. *Journal of Accounting Research* 29 (2): 397–417.

Rajan, R., and L. Zingales. 1995. What do we know about capital structure? Some evidence from international data. *The Journal of Finance* 50: 1421–1460.

Richardson, S., S. H. Teoh, and P. Wysocki. 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.

Rogers, J., and P. Stocken. 2005. Credibility of management forecasts. *The Accounting Review* 80: 1125–1162.

Skinner, D., and R. Sloan. 2002. Earnings surprises, growth expectations, and stock returns: Don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7: 289–312.

Stickel, S. 1991. Common stock returns surrounding earnings forecast revisions: More puzzling evidence. *The Accounting Review* 66 (2): 402–416.

Williams, P. 1996. The relation between a prior earnings forecast by management and analyst response to a current management forecast. *The Accounting Review* 71 (1): 103–113.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1647–1668

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1647

The Effect of Magnitude of Audit Difference and Prior Client Concessions on Negotiations of Proposed Adjustments

Richard C. Hatfield
Richard W. Houston
The University of Alabama
Chad M. Stefaniak
Oklahoma State University
Spencer Usrey
Mississippi State University

ABSTRACT: This study reports the result of an experiment examining two aspects of the audit context that auditors likely do not suspect can influence audited account balances: the magnitude of an audit difference and the presence of a prior client concession. Negotiation theory shows that negotiators' initial positions (e.g., clients' unaudited balances) as well as feelings of reciprocity created by prior negotiations serve to create expectations for the current negotiation and, in turn, affect the outcomes of such negotiations. Our results show that the magnitude of an audit difference involving an estimate (i.e., difference between client's account balance and the auditor's independent estimate) as well as the presence of a prior client concession influence auditors' negotiation expectations. Specifically, auditors proposed smaller adjustments when the magnitude of the audit difference was high and when the client conceded on an audit issue prior to resolving the difference in estimates. These manipulations similarly influence the negotiated outcome, and this influence is fully mediated by the auditor's initial negotiation position.

Keywords: *auditor negotiation; audit adjustments; accounting estimates; audit differences; client concessions.*

We acknowledge Chris Agoglia, Bruce Barrett, Bob Braun, Michelle Diaz, Anna Gold, Steve Kachelmeier (senior editor), Kathryn Kadous (editor), Robert Knechel, Bill Messier, Hun-Tong Tan, and two anonymous reviewers; workshop participants at Florida State University, Louisiana State University, and the University of Wisconsin–Milwaukee; and attendees at the 2008 ISAR Conference, the 2008 Accounting, Behavior and Organizations Conference, and the 2009 University of Florida International Conference on Assurance and Governance for their helpful comments. We also thank the Culverhouse College of Commerce and Culverhouse School of Accountancy for financial support. Finally, we thank all of the firms that provided participants for their support.

Editor's note: Accepted by Kathryn Kadous.

Submitted: July 2008
Accepted: January 2010
Published Online: August 2010

Data Availability: *The data used in this study are available upon request.*

I. INTRODUCTION

Negotiations between auditors and clients are quite common (Gibbins et al. 2001), and they directly affect values reflected in the audited financial statements (e.g., see Brown and Wright 2008). Consequently, the financial statements can be viewed as a product of the joint effort of clients and auditors (Antle and Nalebuff 1991), making these negotiations a crucial link between the quality of the audit and the quality of the financial statements. In this study, we examine two common characteristics of the auditor's negotiation environment: the magnitude of the difference between an unaudited account balance and the auditor's independent estimate of that account balance (i.e., the magnitude of an audit difference) and prior concessions by client management. While auditors likely do not expect these two characteristics of the negotiation context to result in different balances in the audited financial statements, prior negotiation literature implies that these two characteristics likely will significantly influence auditor-client negotiation outcomes.

Prior negotiation research demonstrates that the magnitude of the opening move (i.e., initial negotiation position) by a negotiator's counterpart is important in determining how the negotiator will set his/her negotiation goals, which in turn significantly influence the outcome of the ensuing negotiation (e.g., Rubin and DiMatteo 1972). Based on this perspective, we assert that the client's reported account balances in the unaudited financial statements (e.g., a significant accounting estimate) acts as, either purposefully or not, their "opening move" with regard to subsequent discussions regarding those accounts, although it is unlikely that auditors view it this way. Based on relevant negotiation theory, we predict that auditors' negotiation goals and limits, as well as their initial negotiation positions, will be farther from the audit team's preliminary estimate when the magnitude of the difference between the client's account balance and the auditor's independent estimate is larger.

In addition, audits often include multiple issues that clients and auditors must resolve, and about half the time they resolve these issues sequentially rather than simultaneously (Gibbins et al. 2001). From an auditing perspective, resolution of a subjective issue, such as an estimate, should not be influenced by whether the auditor and client have previously resolved a prior issue (e.g., a known misstatement in the financial statements). However, the social norm of reciprocity has been found to influence the negotiation process as well as subsequent outcomes (e.g., Pruitt and Carnevale 1993; Sanchez et al. 2007; Hatfield et al. 2008). Consistent with reciprocity, when a client agrees to an auditor's request to correct a known misstatement in the financial statements (i.e., a prior client concession), the auditor is likely to feel the need to reciprocate the client's concession in the subsequent negotiation of a more subjective/contentious issue.

To investigate these issues, we conduct a computer-based experiment in which audit partner and manager participants resolve two audit issues with a hypothetical client. We manipulate two independent variables, each at two levels. The first independent variable is the relative magnitude of an audit difference (allowance for obsolete inventory). The audit team's independent estimate is held constant while the client's unaudited account balance varies, resulting in high- or low-difference conditions. The second independent variable is the presence of a prior client concession (i.e., whether a significant, but quantitatively immaterial, known misstatement was resolved prior to or after the inventory issue). We measured various aspects of the auditor's expectations with respect to an ensuing negotiation with the client and asked participants to engage in a simulated negotiation concerning the inventory issue.

Our results are consistent with our predictions. Auditors' pre-negotiation positions differed more from the audit team's independent estimate when the magnitude of the audit difference is high. Additionally, participants' initial negotiation positions reflected smaller adjusted account

balances in the presence of a prior client concession. Both independent variables influenced the negotiated outcome and, consistent with expectations, this effect was fully mediated by the participants' initial negotiation position (i.e., initial offer in the negotiation). Additional analyses show that more auditors take initial negotiation positions and negotiate account balances *materially below* the audit team's estimate when the magnitude of the audit difference is high, and in the presence of a prior client concession. These results add to the growing body of literature demonstrating that various aspects of the audit negotiation context have a direct impact on audit and financial statement quality.

Kayla Gillan (then a PCAOB board member) stated in a speech concerning PCAOB inspections that human nature (i.e., unconscious biases), rather than ethical issues, influences auditor judgments and contributes to audit deficiencies (Gillan 2007). She asserted that making auditors aware of these unintended influences is essential to minimizing their effect on auditor judgments. While such biases may be difficult to eliminate (e.g., Bazerman et al. 2002), our results should be of interest to researchers as they consider how to improve negotiated audit outcomes and practitioners as they develop negotiation training.[1] For example, can explicit consideration of the magnitude of an audit difference or the presence of a prior client concession mitigate the biases demonstrated in our study?

Our mediation results show that the auditor's initial negotiation position is a critical factor in auditor/client negotiations. This result validates prior auditor/client negotiation literature that relies on pre-negotiation measures such as first counter-offer as primary dependent variables (e.g., Bame-Aldred and Kida 2007; Hatfield et al. 2008; Trotman et al. 2009). Finally, with the increased importance placed on estimates involving fair-value accounting, issues such as those addressed in this study (i.e., discussions between auditors and clients over appropriate estimates) will be encountered increasingly in practice.

Section II discusses auditing standards related to accounting estimates and summarizes negotiation research relevant to our study as the basis for developing hypotheses. Section III describes the research design, and Section IV reports results of our analyses. Section V concludes with a summary of the findings and their implications for auditors and audit research.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

Auditing Accounting Estimates

SAS No. 57, *Auditing Accounting Estimates* (AICPA 1989), provides guidance related to auditing clients' accounting estimates.[2] Management is responsible for generating estimates included in the financial statements based on its consideration of relevant factors; ultimately these estimates are based on management's judgment. Auditors are responsible for determining the reasonableness of management's estimates. SAS No. 57 (AICPA 1989, AU 342.04) recognizes that:

> Even when management's estimation process involves competent personnel using relevant and reliable data, there is potential for bias in the subjective factors. Accordingly, when planning and performing procedures to evaluate accounting estimates, the auditor should consider, with an attitude of professional skepticism, both the subjective and objective factors.

---

[1] While auditor-client negotiations are critically important to financial statement values, only 28 percent of auditor participants we surveyed receive formal negotiation training. To date, only Trotman et al. (2005) have examined how to improve auditor negotiations, finding that pre-negotiation consideration of the client's position improves negotiated audit outcomes.

[2] SAS No. 57 specifically cites "net realizable value of inventories where future selling prices and future costs are involved," the subjective accounting issue we examine, as an example of a potentially significant accounting estimate.

Professional guidance requires auditors to gain an understanding of how management develops estimates and evaluate the reasonableness of estimates based on a combination of approaches, which can include developing an independent expectation. SAS No. 107 (AICPA 2006a) states that auditors should treat differences between the auditor's independent estimate and management's estimate as likely misstatements, even if management's estimate is not explicitly deemed unreasonable. Consequently, the magnitude of the initial difference between management's estimate and the auditor's independent estimate should not affect the amount that the auditor ultimately believes should be included in the financial statements.

Auditor-Client Negotiation Literature

While management bears ultimate responsibility for the financial statements, the financial statements can be viewed as a joint effort of both management and the auditor, with account balances often depending on the auditor's approach to negotiating audit differences (Antle and Nalebuff 1991). Gibbins et al. (2001) report that 67 percent of audit partners surveyed enter into negotiations with more than half of their clients, and that all partners negotiate with at least some clients. Therefore, the link between audit quality and the quality of the financial statements clearly is dependent on the nature of negotiations between auditors and clients to resolve proposed audit differences.

Research concerning the disposition of proposed audit adjustments initially did not consider explicitly the underlying negotiation process. Such research was concerned more with characteristics of the proposed adjustment or the audit engagement that influence the likelihood of waiving proposed adjustments. This research found that adjustments are more likely to be waived when, for example, they are immaterial or income-increasing (Wright and Wright 1997; Nelson et al. 2002); the salience of factors related to qualitative materiality is low (Ng and Tan 2007); the client is small or internal controls are strong (Joe et al. 2008); or the risk of waiving the adjustment is low (Braun 2001).

Survey research provides a practical and descriptive perspective on auditor-client negotiations that allows researchers to develop an understanding of the experiences and incentives affecting the negotiating parties (e.g., Gibbins et al. 2001; Gibbins et al. 2005; Gibbins et al. 2007). A growing body of experimental research examines the auditor-client negotiation process. This research finds that negotiated outcomes and aspects of the negotiation process are influenced by factors such as engagement risk (Johnstone et al. 2002); audit committee effectiveness or existence of authoritative guidance (Ng and Tan 2003); auditor experience (Brown and Johnstone 2005); and process variables such as altering the auditor's strategy/approach to the negotiation process (Trotman et al. 2005; Sanchez et al. 2007; Hatfield et al. 2008; Tan and Trotman 2009).

While we extend the research on contextual factors that affect auditor-client negotiations, we also complement and extend prior studies that measure potential negotiation outcomes through the use of pre-negotiation measures (e.g., Bame-Aldred and Kida 2007; Hatfield et al. 2008; Trotman et al. 2009).[3] These studies are based on the premise that pre-negotiation measures (e.g., the auditor's initial negotiation position, negotiation goal, and negotiation limit) are predictive of the negotiated outcomes. However, this premise has not been established in accounting/auditing where environmental characteristics such as regulations and professional concerns can moderate this link.

[3] There are three pre-negotiation measures that comprise the negotiation range. The negotiator's *limit* represents the lowest bound (in the audit context, the lowest amount allowed without qualifying the opinion). The *initial negotiation position* is the negotiator's starting point in the negotiation and represents the upper bound of the range. The negotiator's *goal* is a pre-negotiation measure of the amount the negotiator hopes or expects to achieve as an outcome.

Accordingly, our study elicits these pre-negotiation measures in a computer-simulated negotiation to provide evidence on the effects of our independent variables. We also provide a measurable link between these pre-negotiation measures and negotiation outcomes.

Unaudited Financial Statement Balances: Client's Opening Move?

When viewed through a negotiation lens, the client's unaudited account balances represent the opening move (i.e., initial negotiation position) in any subsequent negotiations.[4] For example, suppose that the auditor develops an independent estimate for the allowance for obsolete inventory of $14 million. Although the magnitude of any audit adjustment(s) should be based on the auditor's estimate and should not be affected by the magnitude of the difference between the auditor's estimate and the client's unaudited balance, negotiation research implies that the audited balance will be influenced by the magnitude of the client's balance (e.g., an unaudited balance of $8 million versus an unaudited allowance of $11.2 million).

Pruitt (1998) finds that, in negotiation settings, one party's initial negotiation position (or opening move) is a key determinant of the other party's actions in subsequent negotiations. For example, Pruitt and Drews (1969) show that the initial position accounts for 67 percent of the variance in subsequent demands. Therefore, unaudited client account balances (i.e., the client's initial positions) are predictive of their posturing regarding subsequent discussions about those amounts (including levels of concession and negotiation goals). Negotiators often alter their negotiation behavior based on their counterpart's initial position in an effort to make an ultimate resolution more likely (Rubin and Brown 1975). For example, Rubin and DiMatteo (1972) find that, when a counterpart's initial demands are high, negotiators tend to lower their own expectations regarding the outcome of the negotiation and, thus, adopt negotiation positions closer to their counterpart's.

Consider our audit example regarding the allowance for obsolete inventory, where the auditor's independent estimate is $14 million. Negotiation research suggests that an unaudited client account balance of $8 million can cause the auditor to adopt negotiation goals and limits closer to the client's to make resolution more likely. That is, getting the client to move $6 million from its unaudited balance could be perceived as quite difficult, causing the auditor to lower his/her negotiation goals and limits. However, if the unaudited balance is $11.2 million, getting the client to move $2.8 million could be seen as less difficult, resulting in the auditor maintaining higher negotiation goals and limits. Specifically, we expect auditors to reduce their aspiration levels to a greater degree when the difference between their preferred outcome and the client's initial position is larger (i.e., when the magnitude of the initial audit difference is larger; Pruitt and Carnevale 1993). This yields the following hypotheses, stated in the alternative form:

H1a: The magnitude of the auditor's goals and limits in a negotiation will be inversely related to the magnitude of the initial audit difference.

H1b: The magnitude of the auditor's initial position in a negotiation will be inversely related to the magnitude of the initial audit difference.

We note that our predictions for magnitude of the audit difference (Hypothesis 1a and 1b) are consistent with anchoring theory, which would suggest that auditors will insufficiently revise their negotiation positions from the anchor provided by the client (i.e., unaudited account balance). Negotiation research has considered the influence of anchoring in negotiation settings (e.g., Ga-

[4] The client's balance is not necessarily meant by the client to be nor interpreted by the auditor as an opening, strategic move in an ensuing negotiation. However, once it becomes apparent that a negotiation will take place, the client's initial account balance *acts as* an initial negotiation position.

linsky and Mussweiler 2001).[5] However, research in other domains indicates that our setting is not conducive to anchoring. For example, Cotter and Henley (2008) found that only inexperienced negotiators were susceptible to anchoring; our participants are highly experienced auditors (primarily senior managers and partners). Perhaps the key characteristic of our setting that mitigates anchoring is the existence of an alternative to the potential anchor. Research finds that the existence of alternatives to anchors (i.e., independent estimates and/or anchor-inconsistent information) reduces or eliminates anchoring effects (e.g., Lord et al. 1984; Wansink et al. 1998; Kristensen and Garling 2000). Because the audit team's independent estimate serves as a separate, inconsistent alternative to the client's unaudited balance, anchoring could work *against* our expectations related to audit difference magnitude, which are based on negotiation theory.

Multiple Negotiation Issues

Prior negotiation literature finds that outcomes differ when multiple issues are negotiated sequentially versus simultaneously (e.g., Rubin and Brown 1975). Generally, research finds greater integration (i.e., win-win) outcomes with simultaneous discussions. However, the audit setting is unique in that issues can arise at different times during the course of the audit, resulting in a greater likelihood of sequential issue resolution. This characteristic is likely magnified by shorter filing periods, which have resulted in increased interim procedures (i.e., pushing toward a more continuous audit approach; e.g., Lambert et al. 2009). In these settings, auditors and clients often resolve issues without knowing whether there will be additional audit issues to resolve; consequently, while audits likely involve sequential issue resolution, many issues are resolved as if they are the only issue requiring resolution.[6]

Audit issues can be objective or subjective and material or immaterial. Subjective items, such as estimates, represent issues for which the auditor and client can conceivably negotiate over the amount because there is no clearly correct amount. Objective items, such as miscalculations, are amounts identified by the auditor that are clearly incorrect in the unaudited financial statements. For objective items, the amount would either have to be adjusted in the financial statements or waived by the auditor (if immaterial) and reported to the audit committee. That is, the correct amount is not in dispute—the question is whether it should be adjusted.

This study focuses on the resolution of a material, subjective audit difference. Negotiations over such issues are at least somewhat contentious (Wright and Wright 1997) because auditors and clients are likely to differ in their perceptions of what the "correct" amount should be. The disposition of objective items tends to be less contentious because the correct treatment of the item is clear and auditors are better able to convince their clients to make adjustments related to objective items than subjective items (Wright and Wright 1997). Thus, if a significant objective item is discovered during an audit, auditors are highly likely to insist that the item be corrected.[7] In such settings, clients are likely to readily concede to auditors' requests to make adjustments (Wright and Wright 1997).

---

[5] Much of this research does not provide alternatives to the manipulated anchor and primarily examines the "first mover's" advantage, which occurs because the first negotiator to make an offer potentially sets the range for the ensuing negotiation. For example, when the sticker price on a car is $30,000, the buyer is anchored to that price and will base the counter-offer on the sticker price. This is different from the anchoring literature in accounting where new information is discounted in the presence of an anchor (e.g., Ashton and Ashton 1988). Research shows that first mover's advantage is mitigated when participants explicitly consider anchor-inconsistent knowledge (Mussweiler et al. 2000) or features inconsistent with the anchor (Chapman and Johnson 1999).

[6] This assertion assumes that auditors do not hold onto all such issues and resolve them at the end of the audit. Discussions with auditors, as well as additional survey evidence we collected, suggest that auditors are very unlikely to wait to resolve these issues. For example, in a scenario similar to that used in the experimental instrument, 87.9 percent of auditors we surveyed would resolve a significant audit issue when it occurred rather than hold it until later.

[7] SAS No. 107 (AICPA 2006a) essentially requires that auditors make this type of adjustment (i.e., a significant known misstatement). Specifically, Paragraph 45 states that "the auditor should request management to record the adjustment needed to correct all known misstatements … other than those that the auditor believes are trivial."

The social norm of reciprocity, as applied to negotiation, asserts that, when one negotiating party makes a concession, the other party is expected to, and tends to, reciprocate the concession (e.g., Pruitt and Carnevale 1993; Sanchez et al. 2007; Hatfield et al. 2008). Reciprocity holds even in situations in which negotiators would not have conceded had it not been for the presence of the perceived need to reciprocate. In fact, people often take actions that they normally would dislike in the presence of a perceived need to reciprocate (Cialdini 2001). Further, relatively small concessions can result in relatively large reciprocations (Cialdini and Ascani 1976).

Consequently, an auditor's reaction to a client's agreement to record an objective adjustment as if it is a concession could create an environment in which the auditor feels the need to reciprocate the client's concession when negotiating other audit issues (e.g., a material subjective item). It is important to note that the influence of reciprocity will be manifested as a social need to concede in the negotiation; we do not expect it to influence pre-negotiation aspirations (e.g., goals and limits). Thus, we expect, based on reciprocity theory, that a prior client concession on the objective issue creates social pressure for the auditor to reduce the magnitude of his/her initial negotiation position (i.e., concede on his/her initial counter-offer), yielding the following hypothesis:

H2: In the presence of a prior client concession to the auditor's request to record a significant objective item, the auditor's initial negotiation position for a subsequent material subjective issue will be lower than in the absence of this client concession.

The above discussion proposes that the magnitude of the audit difference for the subjective issue, as well as a prior client concession, will influence the auditor's initial position in the ensuing negotiation over the subjective issue. Recall that a general conclusion in negotiation research is that higher initial positions result in higher negotiation outcomes (e.g., Rubin and Brown 1975). We test this relationship because unique aspects of the audit setting could influence this relationship. For example, Bame-Aldred and Kida (2007) and Ng and Tan (2003) find that auditors are less likely than other negotiators to make concessions during the negotiation process, most likely because of professional concerns that cause auditors to perceive that it is inappropriate to propose one value in an effort to achieve a different value. Thus, in an auditing context, the auditor's initial position is likely to be a particularly predictive value, possibly even the final negotiated outcome (i.e., no concession). Consequently, we expect that the characteristics of the audit context examined in this study will affect the negotiated outcome and that this effect will be caused, in large part, by the auditor's initial position. This leads to the following mediation hypothesis:

H3: The magnitude of the audit difference as well as the occurrence of a prior client concession will affect the negotiation outcome through their effect on the auditor's initial negotiation position.

III. METHOD

Task and Procedure

The 102 participants for which we report results include 40 partners, 60 managers, and two seniors from public accounting firms ranging in size from regional to Big 4. On average, participants had 14.1 years of experience (S.D. = 8.4), and self-assessed their level of experience resolving audit differences with clients as extensive (mean = 9.6, S.D. = 1.9, with 1 = No Experience and 11 = Extensive Experience). Seventy-four participants (73 percent) are male and 28 female. Participants took an average of approximately 18 minutes to complete the case.[8]

[8] We include the two seniors in all analyses because both indicated that they had experience negotiating audit issues. Patterns of cell means and statistical inferences are the same as those reported if these two participants are excluded

We distributed and collected cases with assistance from contact partners at public accounting firms, primarily in the southeastern United States. Each partner provided us with the number of partner and manager participants whom they thought would complete the case. We sent a corresponding number of USB drives containing the case to the respective offices.[9] Completed cases were returned to the contact partner, who returned the USB drives to us. The four versions of the case were randomly distributed to participants and contact partners did not know which version of the case was included on any given USB drive. We received 108 cases and excluded six participants from the analyses, yielding a sample of 102 used in the analyses.[10]

The Case

Each participant received a cover letter and a USB drive containing the case materials. We asked participants to assume the role of manager on the audit of the 2006 financial statements for Wareham Electronics, a publicly traded (on the New York Stock Exchange) manufacturer of electronic products. Participants then viewed the unaudited 2006 and audited 2005 balance sheet and income statement. Participants could choose to review the financial statements at virtually any time while completing the case (see the Appendix for details and examples of screens from the case).

Participants then learned of the existence of an audit issue (either the subjective inventory valuation issue or a quantitatively immaterial objective issue, depending on the condition). For each issue, participants were notified that Wareham's controller requested a meeting to discuss it; to help the manager prepare for the meeting, the audit team prepared a memo to summarize the issue. It is important to note that, when participants were addressing the first audit issue, they were unaware that a second issue would arise. After proceeding to the second audit issue, participants could not revise their responses to the first audit issue and, after completing the second audit issue, participants proceeded directly to the manipulation check questions and could not alter any previous responses.

Objective Audit Issue

The objective audit difference concerned a professional services invoice received in early 2007 that was related to outsourced internal audit services provided in 2006. The amount, $711,000, should have been accrued as an expense in 2006 (the year in which services were rendered), but because of a clerical error by an accounts payable clerk it was recorded as an expense in 2007. The magnitude of the issue, $711,000, was intended to be considered immaterial,

---

from the analyses. Further, because Trotman et al. (2009) and McCracken et al. (2008) suggest differences between partners' and managers' negotiation abilities/strategies, we consider whether experience affects our results. Two measures of participant experience, position (i.e., partner/manager) and years of audit experience, while correlated with each other, are not correlated with any of the other measures. In addition, we included position and years of experience as independent variables in supplemental analyses to test whether they affected the results; there was no evidence that either experience measure was related to any of the dependent measures, nor were the statistical inferences involving the manipulated variables different.

[9] We sent extra USB drives to each contact partner, some of which may not have been distributed; consequently we do not have an exact response rate. However, at least 90 percent of the case materials distributed was completed.

[10] Two participants failed at least one manipulation check and three took less than two minutes to complete the case (one participant did both). We considered two minutes as unreasonable given the average of over 18 minutes to complete the case and that, aside from these three participants, none spent less than five minutes to complete the case. Of these four participants, three provided enough data for analyses; patterns of cell means and statistical inferences are the same if they are included in any analyses. Additionally, we omit two participants who provided materiality assessments (the covariate in our analyses) identified as extreme outliers (i.e., more than two standard deviations from the mean) based on a box plot of responses. If these participants' responses are included in the analyses, patterns of cell means and statistical inferences are the same as those reported.

but significant.[11] We intended to create a scenario in which participants would be inclined to make the adjustment (to be consistent with SAS No. 107; AICPA 2006a).

Participants were provided with the journal entry that would be necessary to correct the error and then were asked to indicate their preference for whether the adjustment should be recorded (with higher numbers indicating a greater preference for adjustment). To ensure that all participants knew that the auditor recommended the adjustment and that the client conceded, when this issue was presented first, they were told that prior to the meeting with Wareham's controller, "you and the audit engagement partner ... agreed that your firm should recommend that Wareham correct this error," that Wareham's controller agreed to make the correcting entry, and that the financial statements were adjusted to reflect the effects of the journal entry. Overall, only 18 percent of participants (18 participants) responded below the scale midpoint.[12]

Subjective Audit Issue

The subjective issue involved the estimate for the allowance for obsolete inventory (see the Appendix for a detailed description). Because a competitor introduced a technologically superior and less expensive product, Wareham needed to increase its existing allowance. The audit team developed an independent estimate ($14 million) that was materially larger than the client's allowance, substantiated by a memo from the audit team. Participants provided expectations regarding the upcoming discussions (i.e., their own goals and limits and their perceptions of the client's goals and limits) and then completed a simulated negotiation with the client.

In the ensuing negotiation, participants were reminded of the magnitude of the audit team's estimate and Wareham's balance and were asked to make one of three choices: (1) accept Wareham's balance; (2) insist that Wareham adjust the balance to $14 million; or (3) propose an adjustment somewhere between the two amounts. If participants chose option 1, they were asked to confirm that they decided to waive the adjustment; they then moved to either the second audit issue or the manipulation check questions, depending upon the condition. If participants chose option 2 or 3, they were told that Wareham responded by offering to increase its account balance.

Each negotiation round started with participants being offered the same three options as on the first negotiation screen (with option 2 being equal to the auditor's previous proposed amount). When a participant contended (i.e., did not move from the balance stated in the prior round), we programmed Wareham's response as an increase of 10 percent of the difference between the auditor's and Wareham's balance. If the auditor contended for three rounds, the negotiation ended, and the auditor required an adjusting entry based on the their balance.[13]

As participants proposed account balances, they viewed the adjusting journal entry that would result. If the auditor proposed a new adjustment (option 3 above), then the client's response was programmed based on prior negotiation research finding that concessions by one party are likely to be matched by the other party (Pruitt 1998). In our case, the client engages in tracking (or concession matching) behavior, a common negotiation strategy in which a negotiator makes con-

---

[11] Arguably, the most commonly used quantitative measure of materiality is 5 percent of net income (e.g., SEC 1999). Based on this measure, quantitative materiality in our case is approximately $1.8 million. We define significant as large enough to be added to a list of unadjusted audit differences that are aggregated to assess whether their cumulative effect is material. Additionally, significant items, if waived, would need to be reported to the audit committee (SAS No. 114; AICPA 2006b).

[12] Most of these 18 participants were near the scale midpoint of 6 (i.e., between 3 and 5), suggesting that these participants only mildly agreed with the statement that the item could be waived. More participants in the objective item first condition fell below the scale midpoint (i.e., would prefer to waive the adjustment). This is consistent with the idea that participants who received the objective item second and had already made concessions in the negotiation involving the subjective issue were less likely to concede further.

[13] While negotiating, participants were not aware of when negotiations would end (i.e., the maximum number of rounds). In addition, during negotiations, the auditor's new proposed balance was included in all subsequent negotiation screens.

cessions similar in amount as their counterparts' (Pruitt 1981).[14] Specifically, Wareham increased its previous balance by either 110 or 90 percent of the amount that the auditor moved from his/her previous offer (we alternated between 110 and 90 percent to simulate a negotiation in which client counter-offers did not obviously match the auditor's movement).

To ensure that the case took a reasonable amount of time to complete, the ensuing negotiation could last up to five rounds and would end at any time if the difference between the auditor and client balances was less than $200,000. At the end of the negotiation, participants were shown the journal entry that Wareham would make and were told that the financial statements were adjusted to reflect the effects of this entry (participants had the opportunity to view the adjusted financial statements). Upon resolution of the second audit issue, participants completed manipulation check questions related to the two independent variables, provided their assessment of materiality (based on the unadjusted financial statements), and answered several questions related to the case and themselves (including demographic questions). In our statistical analyses, we include materiality assessments as a covariate because we believe that the extent to which individuals differ with regard to materiality assessments could influence the size of preferred audit adjustments. Consequently, our analyses control for random variation in the dependent variables caused by differences in individuals' materiality assessments.[15]

Independent Variables

To test the hypotheses, we utilize a 2 × 2 between-subjects experimental design that manipulated (1) the magnitude of the audit difference; and (2) whether the client conceded on an objective, immaterial but significant, issue prior to resolving the inventory valuation issue (i.e., prior client concession). We manipulated the magnitude of the audit difference by varying the client's unaudited inventory valuation balance as either $8 million (i.e., high difference) or $11.2 million (low difference). Recall that the audit team estimate was $14 million, so the magnitude of either audit difference (i.e., $6 million or $2.8 million) is always material. We operationalized the second factor, prior client concession, by altering the order of the two issues. Half of the participants resolved the objective issue first and half resolved the subjective inventory valuation issue first.

To validate our assertion that these issues are addressed sequentially in practice, in a supplemental data collection (detailed in the "Additional Experimentation" section) we presented participants with a scenario identical to the objective issue and asked them when they would resolve it (1 = Definitely resolve now; 11 = Definitely resolve at end of audit). The mean response is 2.89; of the 33 responses, 15 equal 1 and 29 (87.9 percent) are less than 5, which we interpret as strong evidence that auditors would resolve this issue sequentially (i.e., at the time it arose).

Dependent Variables

The primary dependent variables used to test the hypotheses are those elicited during consideration of the subjective inventory issue, including the ensuing negotiation. Specifically, we obtained pre-negotiation amounts by asking participants to provide their negotiation goal and lowest

---

[14] Tracking, or concession matching, is a strategy by which a negotiator will concede a similar amount as his/her counterpart. For example, when a seller comes down $10,000 on the value of a home, a buyer using the matching strategy likely will increase their last offer by around $10,000.

[15] While a covariate is best measured prior to any manipulations to ensure that the manipulations do not influence the covariate, we measured materiality assessments at the end of the case because we were concerned that such a question might influence participants' negotiation behavior. Because materiality assessments could have been associated with the independent variables, we examined the correlations between the independent variables and materiality assessments. Prior to the removal of two extreme outliers (discussed in footnote 10), Prior Concession was marginally correlated with materiality assessments; however, this correlation became insignificant when these observations were removed, suggesting that materiality is an appropriate covariate.

acceptable balance (i.e., negotiation limit) for the allowance account. Such pre-negotiation measures are employed in audit negotiation research as they are indicative of the resulting negotiated outcomes (e.g., Bame-Aldred and Kida 2007; Trotman et al. 2009). At the start of the negotiation process, participants indicated their initial negotiation position, which serves as a pre-negotiation measure and the mediating variable for the test of H3. Finally, the outcome that results at the conclusion of the negotiation is used as the dependent variable for H3.

To provide a supplemental measure of practical significance, we considered, for each independent variable, whether participants' goals, limits, initial negotiation positions, and negotiation outcomes were materially lower than the audit team's estimate. For this analysis, we subtract the appropriate materiality figure from the audit team estimate and determine the frequency with which the respective dependent variables fall below this figure. We first use a common quantitative assessment of materiality, 5 percent of net income (e.g., SEC 1999; Messier et al. 2008). For robustness, we also consider two measures of account level materiality (i.e., tolerable misstatement), 75 and 50 percent of overall materiality.[16] These tests provide evidence indicating how often the manipulated variables result in economically and practically significant differences relative to the audit team's independent estimate.

IV. RESULTS

Magnitude of the Audit Difference (H1a and H1b)

Hypotheses 1a and 1b predict that the magnitude of the auditor's goals and limits as well as their initial negotiation position will be inversely related to the magnitude of the initial audit difference for the inventory valuation issue. Table 1 reports the means for the dependent variables and a summary of results of t-tests that examine the effects of prior concession and magnitude of the audit difference on the dependent variables.[17] Hypothesis 1a predicts that auditors' pre-negotiation positions will be lower when the magnitude of the audit difference is relatively high than when it is relatively low (higher positions reflect more conservative account balances). The auditors' goal when the magnitude of the audit difference was high ($13,283,205) was significantly less than when the magnitude was low ($13,644,021; $p = 0.02$). Further, auditors' limit when the magnitude of the audit difference was high ($12,469,147) was significantly less than when the magnitude was low ($13,163,508; $p < 0.01$), supporting H1a. These results suggest that auditors' aspirations regarding the negotiation outcome are lower when the magnitude of the difference is higher (i.e., reflecting a less conservative account balance). Auditors' initial negotiation positions are also significantly lower when the magnitude of the audit difference is high ($12,929,204) compared to when the magnitude is low ($13,437,903; $p < 0.01$), supporting H1b.

Table 2 reports the results of one-tailed Fisher's Exact Tests that provide additional evidence regarding economic (i.e., practical) significance. Specifically, we examine the proportion of auditors whose initial negotiation position is materially below the audit team's estimate. Consistent with H1b, when the magnitude of the difference is low, only 4.1 percent of participants have an

[16] Our review of two large international accounting firms' audit manuals, conversations with Big 4 audit partners, and various audit texts (e.g., Messier et al. 2008; Puncel 2008) reveal that account balance materiality (i.e., tolerable error) typically is between 50–75 percent of overall materiality (with 50 percent being the most conservative measure).

[17] We first conducted ANCOVAs for the full 2 × 2 design for all dependent variables described in the "Method" section (preceded by a MANCOVA because the dependent variables are highly correlated). We report results for H1 and H2 using t-tests that take the covariate into account by using the mean squared error (i.e., MSE) from the ANCOVA to compute single degree of freedom contrasts. As all hypotheses are directional, associated p-values are one-tailed and, because we base our analyses on ANCOVA, we report adjusted means. The ANCOVA results, not tabulated, are consistent with the results of these tests, and there were no significant interactions in any of the analyses. We also conducted the analyses based on converting our dependent variables to ranked data because there is some evidence that our dependent variables violate the normality assumption. The results based on ranked data yield patterns of cell means and statistical inferences that are the same as those reported in Table 1.

TABLE 1

Hypothesis Tests

Dependent Variables	Magnitude of Difference Mean (Std. Dev.)[a] High n = 53	Magnitude of Difference Mean (Std. Dev.)[a] Low n = 49	Prior Concession Mean (Std. Dev.) Yes n = 52	Prior Concession Mean (Std. Dev.) No n = 50	Statistical Tests[b] t-statistic (p-value) [Hypothesis] Magnitude of Difference	Statistical Tests[b] t-statistic (p-value) [Hypothesis] Prior Concession
Auditor's Negotiation Goal	13,283,205 (1,045,099)	13,644,021 (587,311)	13,347,938 (928,609)	13,579,288 (825,837)	2.08 (0.02) [H1a]	1.36 (0.18) [NA]
Auditor's Negotiation Limit	12,469,147 (1,104,601)	13,163,508 (868,177)	12,758,763 (1,001,422)	12,873,892 (1,158,092)	3.56 (<0.01) [H1a]	0.60 (0.55) [NA]
Auditor's Initial Negotiation Position	12,929,204 (1,017,842)	13,437,903 (644,677)	13,040,122 (928,490)	13,326,985 (862,288)	2.91 (<0.01) [H1b]	1.67 (0.049) [H2]
Negotiation Outcome	12,740,492 (992,340)	13,196,996 (742,461)	12,787,392 (1,001,904)	13,150,096 (1,008,523)	2.34 (0.01) [H3]	1.89 (0.03) [H3]

[a] All reported means are adjusted means from an ANCOVA with Magnitude of the Audit Difference and Prior Concession as the independent variables and participants' materiality assessments as the covariate.

[b] Because all hypotheses are directional, the reported p-values for hypothesis tests are one-tailed. The results are reported using t-tests following an ANCOVA that use the MSE from the ANCOVA to compute a single degree of freedom contrast.

TABLE 2

Tests of Economic Significance
Significance Levels for Fisher's Exact Tests[a]

Dependent Variables	Materiality Magnitude of Difference	Materiality Prior Concession	75% Materiality Magnitude of Difference	75% Materiality Prior Concession	50% Materiality Magnitude of Difference	50% Materiality Prior Concession
Auditor's Negotiation Goal	<0.01	0.76	<0.01	0.61	0.03	0.09
Auditor's Negotiation Limit	<0.01	1.00	<0.01	1.00	<0.01	0.31
Auditor's Initial Negotiation Position	<0.01	0.11	0.03	0.05	<0.01	0.06
Negotiation Outcome	<0.01	0.10	0.09	0.03	0.04	0.02

[a] Numbers in table are p-values based on Fisher's Exact Test. All p-values one-tailed (consistent with hypothesis tests) except for the tests of Prior Concession for Auditor's Negotiation Goal and Auditor's Negotiation Limit (which are two-tailed), because we do not provide related hypotheses. Significant p-values indicate that participants' goals, limits, initial negotiation positions, as well as negotiated outcomes were more likely to be materially below the audit team's estimate when the magnitude of the audit difference is high or if there is a prior client concession.

initial negotiation position materially below the audit team's estimate, as compared to 30.2 percent when the magnitude of the difference is high ($p < 0.01$). Results are similar ($p = 0.03$ and $p < 0.01$) when using the endpoints of the range for tolerable misstatement as the basis of comparison (i.e., 75 and 50 percent of materiality, respectively). Consistent with H1a, we find similar results for the auditor's goal (limit). Specifically, for goal (limit), the low-magnitude audit difference results in 2.0 percent (16.0 percent) of participants providing values materially below the audit team estimate, as compared to 18.9 percent (51.9 percent) when the magnitude of the difference is high (all p-values < 0.01). Inferences are the same when considering either measure of tolerable misstatement. These tests provide evidence that auditors are more likely to move materially toward the client's unaudited balance when the magnitude of the audit difference is high.

Prior Concession (H2)

Hypothesis 2 predicts that when a client makes a concession prior to the negotiation, the auditor will feel the need to reciprocate the concession, resulting in a lower initial negotiation position. Results reported in Table 1 indicate that, in the presence of a prior client concession, auditors' mean initial negotiation position ($13,040,122) is lower than in the absence of a prior concession ($13,326,985; $p < 0.05$), providing support for H2. While the difference in mean responses ($286,863 for initial negotiation position and $362,704 for negotiation outcome) does not rise to the level of materiality or even tolerable misstatement, misstatements of this amount (about 15 percent of materiality) would likely be large enough to be listed in a summary of unadjusted audit differences and, if waived, disclosed to the audit committee (e.g., SAS No. 114; AICPA 2006b).[18]

Further analysis in Table 2 to provide evidence of economic significance shows that 23.1 percent of auditors in the prior concession condition had an initial negotiation position materially below the audit team's estimate, as compared to 12.0 percent in the absence of a concession ($p = 0.11$; one-tailed Fisher's Exact test). When using tolerable misstatement as the basis for comparison, results are stronger ($p = 0.05$ and $p = 0.06$ for 75 percent and 50 percent of materiality, respectively). In addition, statistical inferences are the same for the negotiation outcome ($p = 0.10$, 0.03, and 0.02 for the three levels of materiality, respectively). These tests provide additional evidence that auditors are more likely to move materially toward the client's unaudited balance in the presence of a prior client concession.

Negotiated Outcome: Mediating Role of Initial Demand (H3)

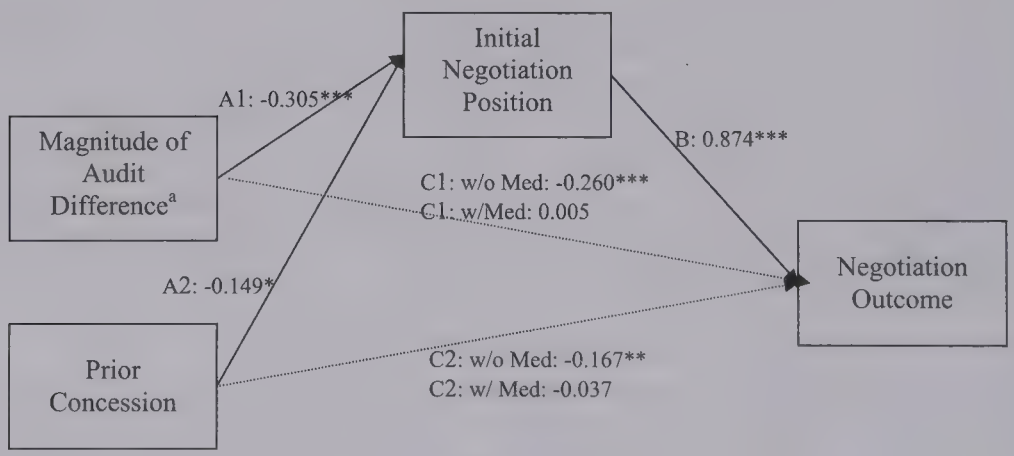
The dependent variables examined in the first two hypotheses are based on prior research demonstrating that negotiators' pre-negotiation positions (goals and limits) as well as their initial negotiation positions are strongly correlated with negotiators' eventual success in the negotiation. That is, higher goals, limits, and initial demands result in higher outcomes for a negotiator (Rubin and Brown 1975). However, to obtain further confidence that the independent variables considered in this study influence the negotiation, we simulate a negotiation, via computer interaction, with the client's controller. As discussed above, the client negotiator is programmed with a matching strategy, which is a predominant negotiation style in this context (Pruitt 1998). Using the outcome of the negotiation as the dependent variable, both magnitude of the audit difference and prior concession have a significant effect on the negotiation outcome (one-tailed p-values = 0.01 and = 0.03, respectively; see Table 1), with the pattern of means and statistical inferences consistent with those reported above.

[18] For example, one large international firm we consulted has guidelines stating that its *Audit Difference Posting Threshold* is 3 to 5 percent of materiality, and another indicated that its threshold is 5 percent of materiality, both well below the differences we observe.

Hypothesis 3 predicts that both manipulations will affect the negotiation outcome and that these effects are attributable to their influence on the auditor's initial negotiation position. We test this prediction, diagramed in Figure 1, with mediation analyses based on Baron and Kenny (1986). The mediating relationship will be demonstrated if three conditions are met: (1) magnitude of audit difference and prior concession influence negotiation outcome (paths C1 and C2 in Figure 1), (2) magnitude of audit difference and prior concession significantly influence initial negotiation position (paths A1 and A2 in Figure 1), and (3) when the initial negotiation position is included in the model, the effect of the manipulated variables become insignificant. The resulting beta coefficients of these regressions and their significance levels are included in the paths of the mediating model in Figure 1.

Results indicate that condition 1 is met, as the coefficients for paths C1 and C2 are significant when the mediator is not included in the model. Further, condition 2 is met, as paths A1 and A2 are significant. Finally, full mediation is indicated, as the effects of both magnitude of audit difference and prior concession (paths C1 and C2) become insignificant when the mediating variable (i.e., initial negotiation position) is added to the model (condition 3) and path B is significant. These results indicate that the effect of our two independent variables on the negotiation outcome is attributable to their influence on auditors' initial negotiation positions (i.e., medi-

FIGURE 1
Mediating Relationship
(Beta Coefficients for Paths)



*, **, *** $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.
Equation (1): Initial Negotiation Position = A1 (magnitude of estimate difference) + A2 (prior concession)
Equation (2): Negotiation Outcome = B (Initial Negotiation Position)
Equation (3): Negotiation Outcome = C1 (magnitude of estimate difference) + C2 (prior concession)
Equation (4): Negotiation Outcome = C1 (magnitude of estimate difference) + C2 (prior concession) + C3 (Initial Negotiation Position)

[a] We coded Magnitude of Audit Difference as 1 if high, and 0 if low. Prior Concession is coded as 1 in the presence of a prior concession, and 0 in its absence. For ease of interpretation, the analyses in this figure do not include Materiality as a control variable; if materiality is included, it is not significant in any of the equations, p-values and Beta coefficients generally are virtually the same as those reported, and full mediation still occurs.

ated paths A1 and A2, and path B). While it is possible that the mechanical nature of the computer tracking strategy drove the mediation finding, a closer examination of the data suggests that the auditors' use of a contending strategy possibly drives the mediation results. Seventy-six percent of participants stick to their initial offer without conceding (same percentage across conditions). This level of contention demonstrates a difference from the audit and psychology negotiation literature and appears to result in a particularly strong link between initial negotiation position and negotiation outcome.[19]

Additional Experimentation

We gathered additional data to rule out potential alternative explanations for each independent variable. First, while we use negotiation theory as a basis for predicting the findings related to the magnitude of the audit difference, a possible alternative explanation is that participants were less confident in the accuracy of the audit team's estimate when it differs more from the client's balance. Consequently, we test whether auditors' confidence in the audit team's estimate differs based on the magnitude of the audit difference. Second, while we base our prediction for prior concession on reciprocity theory, participants could view a prior concession as an indication that the client is an "easy negotiator," and thus start the negotiation closer to the client's balance, which yields the same directional prediction. Accordingly, we test whether auditors' expectations of the difficulty of the negotiation vary based on whether the client has made a prior concession.

Participants in the additional experiment included 33 audit partners and managers from large accounting firms. Compared to the main experiment, participants included a higher proportion of partners (61 percent versus 39 percent in the main experiment; Fisher Exact Test $p = 0.04$) and had more work experience (17.5 years versus 14.1 in the main experiment; $p = 0.04$). However, as in the main experiment, when we control for these measures of experience, statistical inferences are the same as those reported. Participants' level of experience resolving audit differences with clients did not differ significantly across the two experiments (means of 9.5 and 9.6, with 11 representing extensive experience, for the additional and main experiments, respectively; $p = 0.80$). To examine the potential competing explanations for each of the independent variables, we conducted two 1×2 analyses.

Confidence in the Audit Team Estimate

The first analysis varies the magnitude of the audit difference to consider whether larger differences led participants to believe that the audit team's estimate was of lower quality. We provided participants with information almost identical to that in the main experiment; however, we conducted this experiment on paper, precluding us from simulating the negotiation.[20] We measured the auditor's goal, limit, and initial negotiation position. Results reported in Table 3 demonstrate that H1a and H1b are supported in the supplemental data collection. In addition, results not tabulated indicate that statistical inferences from tests of economic significance identical to those reported in Table 2 are the same as for the main experiment.

---

[19] Of the 78 participants who did not move from their initial position, 43 (55 percent) proposed an account balance below the audit team's estimate. The mean and median for these participants' initial positions (and outcome, since they did not move from their initial position) was approximately $12.6 million. It is interesting to note that, on an overall basis, these participants' initial positions did not differ significantly from those who did move from their initial position (mean of $12.9 million; $t = 1.27$, $p = 0.21$). For participants who did move from their initial position, the negotiation lasted an average of 3.42 rounds, with no differences between conditions.

[20] The only difference from the main test of H1 conducted in the primary experiment is that we use $11 million rather than $11.2 million as the initial client balance to rule out the possibility that an estimate of $11.2 was viewed as more precise than $8.0 or $14.0, and thus was given more weight.

TABLE 3

Summary of Results for Additional Experimentation

Dependent Variables	Magnitude of Difference Mean (Std. Dev.) High n = 15	Low n = 18	t-statistic
Auditor's Negotiation Goal	13,136,960	13,630,687	1.52*
	(1,109,215)	(600,245)	H1a
Auditor's Negotiation Limit	10,949,312	12,050,688	2.54***
	(1,659,568)	(884,008)	H1a
Auditor's Initial Negotiation Position	11,788,670	12,791,885	2.38**
	(1,626,968)	(738,534)	H1b
Confidence in Audit Team's Estimate	7.60	7.06	0.88
	(1.24)	(2.10)	

*, **, *** Indicates $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively. The p-value for test of Confidence in Audit Team's Estimate is two-tailed, as we do not make a directional prediction.

All other p-values are one-tailed, consistent with the hypotheses (relevant hypotheses are noted in the table). Results for hypothesis tests are based on one-tailed t-tests following an ANCOVA with Magnitude of the Audit Difference as the independent variable and participants' materiality assessments as the covariate. These t-tests use the MSE from the ANCOVA to compute a single degree of freedom contrast. Because we base our analyses for auditor's negotiation goal, negotiation limit, and initial negotiation position on ANCOVA, means for these variables are adjusted means.

The p-value for Confidence in the Audit Team's Estimate is based on a two-sample t-test. Confidence in the Audit Team's Estimate is measured on an 11-point scale with higher numbers indicating greater confidence. All other variables measured in dollars.

A potential alternative to negotiation theory is that, in our setting, auditors are less confident in the accuracy of their audit team's estimate when the difference from the client's balance is larger. Thus, it could be an explicit consideration of the quality of the information provided to the participant rather than the influence of the client's initial position, as suggested by negotiation theory. We address this alternative explanation by asking participants to respond to the following question after they were presented with the client's balance and the audit team's independent estimate: "How confident are you in the accuracy of your audit team's independent estimate of the allowance for obsolete inventory?" (1 = Not at all confident; 11 = Very confident). Results indicate that confidence in the audit team's estimate is virtually the same across condition (7.60 and 7.06 for high and low magnitude of the audit difference, respectively; two-tailed $p = 0.38$). Consequently, it appears that participants' confidence in the audit team's estimate does not drive the observed difference in negotiation behaviors.[21]

Expected Negotiation Difficulty

Participants could view a prior concession as indicative of the difficulty of subsequent negotiations. That is, auditors could perceive the concession as a signal that the client is an easy negotiator, which would cause auditors to expect easier negotiations to get the desired outcome, and therefore take an initial negotiation position closer to the client's balance. This view provides

[21] In addition, confidence in the estimate did not approach significance as a control variable in analyses including any of the negotiation measures. Confidence in the estimate exceeded the scale midpoint for both conditions ($p < 0.01$), suggesting that auditors were confident in the accuracy of the audit team's estimate.

a similar directional prediction as our argument that the result is attributable to reciprocity.[22] To address this counter-explanation, we conducted a second 1 × 2 analysis by presenting the same 33 auditors with a separate case in which they viewed one of two scenarios intended to replicate the prior concession condition in the main experiment.[23]

In one condition, we described the resolution of an objective issue (the same issue and the same dollar amount as in the main experiment) and then provided a subjective issue; specifically, we told participants that the audit team developed a balance for the allowance for bad debts significantly higher than the client's, and that the participant would meet with the client to resolve the issue (dollar amounts were slightly different from the primary study to ensure that participants viewed this issue as distinct from the inventory issue). In the other condition, we included only the subjective allowance issue. We asked all participants to assess "How do you expect the negotiations with the client to be when you try to resolve this issue?" (1 = Very easy; 11 = Very difficult). Results (not tabulated) demonstrate virtually no difference between perceived difficulty of the negotiation when only the subjective issue was present (mean = 7.44, n = 16) than when it was preceded by the objective issue (mean = 7.53, n = 17; p = 0.87).

V. DISCUSSION AND CONCLUDING REMARKS

This study reports the results of an experiment that considers two aspects of the auditor's negotiation environment that auditors possibly do not suspect will result in different balances in the audited financial statements: the magnitude of the difference between a client's unaudited balance for a subjective, material accounting issue relative to the audit team's independent estimate and the occurrence of a prior concession by management on a significant objective item. We find that both factors affect auditors' negotiation expectations and behavior and, ultimately, the amounts reflected in the financial statements—in fact, the negotiation outcome is fully mediated by auditors' pre-negotiation behavior. We also report results suggesting that these differences are economically significant. Our findings are consistent with the premise from the negotiation literature that a negotiation counterpart's initial position (e.g., a client's unaudited balance) as well as feelings of reciprocity created by prior concessions set expectations for the current negotiation, shaping the outcomes.

These findings highlight important considerations for auditors who resolve audit differences with their clients. The greater a difference between the client's unaudited account balance and the auditor's independent estimate for that account, the less likely it is that the auditor will adjust that account balance close to the auditor's estimate. These results are important because (1) *ceteris paribus*, the auditor's proposed adjustment should be unaffected by the magnitude of the audit difference when the auditor believes its independent estimate is appropriate; and (2) management, in many cases, has incentives to bias estimates in the financial statements, and its unaudited balances (especially in accounts involving subjectivity, such as estimates) likely reflect its incentives to bias. This "non-normative" effect is particularly highlighted by the result that auditors materially alter the point at which they would qualify the financial statements between conditions (i.e., negotiation limit).

Our prior concession results suggest that the order in which multiple issues in a negotiation are addressed also can affect negotiation outcomes. With regard to situations in which issues differ in importance, researchers are divided on whether the more or less important issue should be

[22] While we assert that our result is attributable to reciprocity, we did not elicit from participants a perceived need to reciprocate. Consequently, we cannot directly address that process, representing a limitation of our study.

[23] Because participants in the supplemental data collection responded to both experimental scenarios, we counterbalanced the scenarios and manipulations, resulting in four different versions of the case. Patterns of cell means and statistical inferences are the same regardless of scenario order or differing manipulations in the other scenario.

addressed first (e.g., Rubin and Brown 1975). We demonstrate that, in an audit setting, the ultimate disposition of audit issues can differ based on the order in which they are addressed. Our results imply that auditors should consider, if possible, the order in which they address audit issues.

We also extend audit negotiation research by simulating a negotiation. By doing so, we provide evidence that pre-negotiation measures used in several prior audit studies (e.g., Bame-Aldred and Kida 2007; Hatfield et al. 2008; Trotman et al. 2009) are indicative of negotiation outcomes. While our results also support statements by Ng and Tan (2003) and Bame-Aldred and Kida (2007), who find that auditors are less likely than other negotiators to make concessions during the negotiation process, care should be taken when drawing inferences from these results due to the limitations based on the mechanistic nature of simulated negotiations.

Finally, in a broader sense, we add to research concerning the link between audit quality and financial statement quality. Audit quality has been the focus of several streams of research as well as the object of recent and substantial changes in audit regulation. However, financial statement quality ultimately is contingent upon how the discussions and negotiations between auditors and client management affect adjustments to the financial statements. By demonstrating that the size of audit adjustments are a function of two aspects of negotiation settings to which they should be invariant, we demonstrate that, even when auditors plan and execute quality audits, financial statement quality can ultimately be diminished by contextual features that lead to biases such as those warned against by Gillan (2007).

This study is subject to several limitations, some of which we have already mentioned. For example, prior research demonstrates that, in certain contexts, gender affects negotiation behavior (e.g., Barron 2003; Flynn and Ames 2006), and our sample is primarily comprised of males (73 percent). We find that gender is not significant in any analyses, nor does it interact with any variables. Further, our experimental environment lacks incentives present in typical audit environments (e.g., no consequence to waiving adjustments or disagreeing with clients).

Also, as in most experimental research, we are only able to examine a particular piece of the audit (negotiation) without consideration of other aspects of the audit occurring before and after any negotiations take place. Other aspects of the audit potentially influence the negotiation. For example, Tan (1995) finds that prior involvement with a going-concern decision leads auditors to pay more attention to evidence consistent with that prior finding. Consequently, future research could examine whether prior involvement at other points during the audit (e.g., auditor evaluation of the client's process to develop its estimate, or involvement in developing the audit team's estimate) influences judgments in subsequent negotiations. Further, in practice, it is likely difficult for auditors to ascertain whether an audit difference is large or small, and it is also likely difficult to manage the order in which issues are addressed. However, auditors should be aware that these biases exist and should perhaps receive negotiation training intended to minimize undesirable outcomes. Yet, the auditors we surveyed indicated that the availability of formal negotiation training is limited. We propose that audit firms consider developing formal negotiation training that addresses issues addressed in this study as well as other issues influencing the effectiveness of auditor negotiation behavior.

APPENDIX

With respect to the inventory valuation issue, participants learned that "because of frequent technological advances and associated price reductions in competitors' products, it is important that Wareham constantly monitor its competitors' new product development and pricing to ensure that inventory is properly valued at lower of cost or market." In the audit issue memo prepared by the audit team in anticipation of the meeting with Wareham's controller, participants then were told that, late in 2006, one of Wareham's competitors introduced a new product line that was

technologically superior to and less expensive than Wareham's similar product offerings and that the competitors' products likely would reduce demand and prices for Wareham's products. Because Wareham believed that some of the inventory would be sold at lower prices and some would never be sold, the company increased the balance in its allowance for obsolete inventory from its previous balance of $7.5 million.

The audit team developed an independent estimate for the allowance for obsolete inventory and proposed an adjusting entry based on this estimate. The experimental materials stated that the audit team believed that management's estimate was made in good faith, but that the audit team gathered additional evidence to develop an independent estimate and concluded that the allowance for obsolete inventory should be increased to $14 million. We worded the case materials so that, consistent with SAS No. 57 (AICPA 1989), the difference between Wareham's balance and the auditor's independent estimate should be considered as a likely misstatement. Below are examples of the two of the negotiation screens included in the research case, the second of which was seen if participants chose to propose a different adjustment amount. Note that the amounts for the adjusting entry changed with changes to the proposed balance.[24]

File

Your Meeting with Wareham's Controller

Assume that you are now in the meeting with Wareham's controller to resolve the inventory issue. Recall the following:

Wareham's balance in the allowance for obsolete inventory is **$8,000,000**.

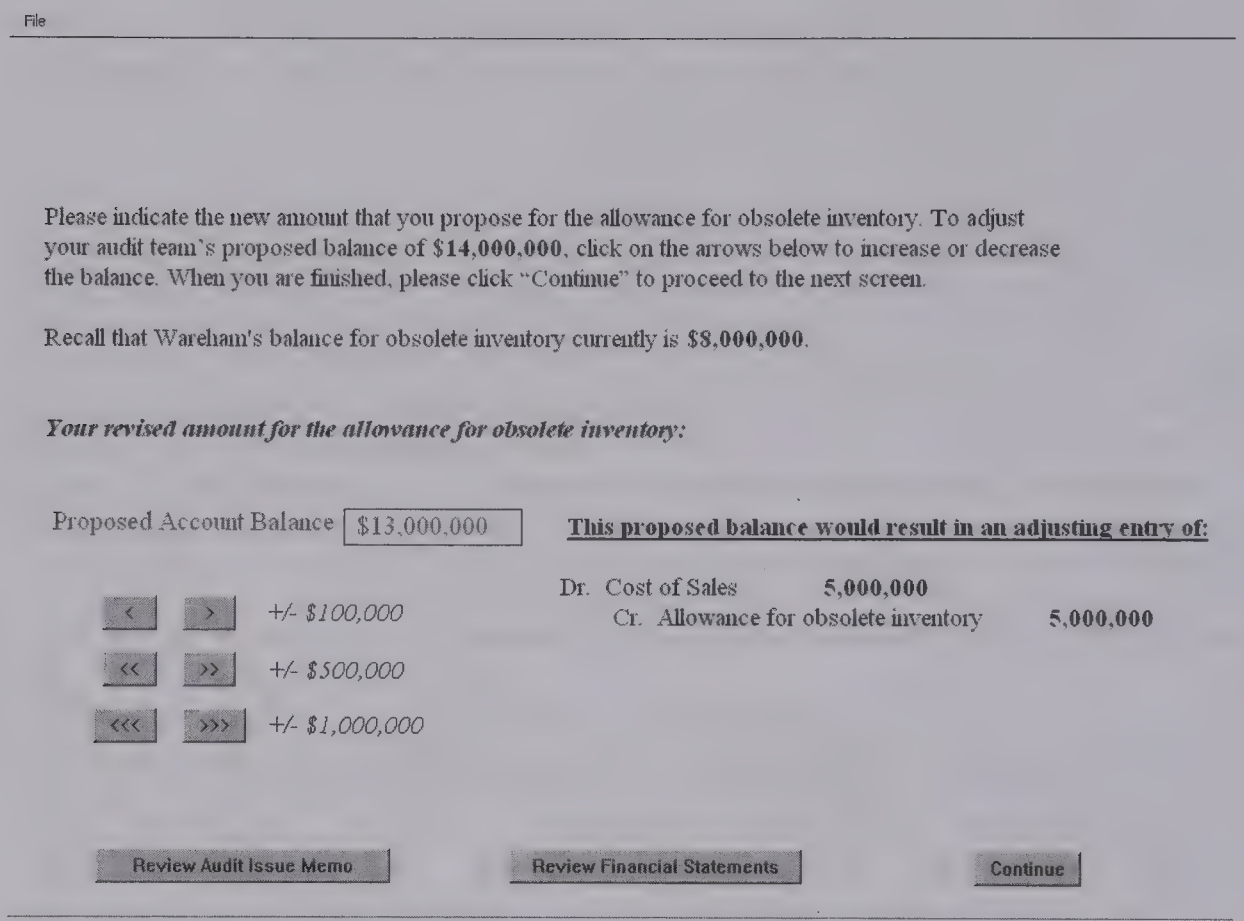
Your team's proposed balance for the allowance for obsolete inventory is **$14,000,000**.

You must now decide how to proceed.
(Please indicate your decision by clicking on the appropriate button below)

1. Accept the amount currently in the financial statements (i.e., **$8,000,000**) Waive Adjustment or,
2. Insist that Wareham adjust the account balance to **$14,000,000** Require Full Adjustment or,
3. Propose an adjustment somewhere between these two amounts Propose New Adjustment

Review Audit Issue Memo Review Financial Statements

[24] Numbers change as the negotiation continues and the client counters with new adjustment amounts. In this example, the auditor counters the client's initial balance with an offer of $13 million.



REFERENCES

American Institute of Certified Public Accountants (AICPA). 1989. *Auditing Accounting Estimates*. Statement on Auditing Standards No. 57. New York, NY: AICPA.

———. 2006a. *Audit Risk and Materiality in Conducting an Audit*. Statement on Auditing Standards No. 107. New York, NY: AICPA.

———. 2006b. *The Auditor's Communication with Those Charged with Governance*. Statement on Auditing Standards No. 114. New York, NY: AICPA.

Antle, R., and B. Nalebuff. 1991. Conservatism and auditor-client negotiations. *Journal of Accounting Research* 29 (Supplement): 31–54.

Ashton, A. H., and R. H. Ashton. 1988. Sequential belief revision in auditing. *The Accounting Review* 63 (4): 623–641.

Bame-Aldred, C. W., and T. Kida. 2007. A comparison of auditor and client negotiation decisions. *Accounting, Organizations and Society* 32 (6): 497–512.

Baron, R., and D. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–1182.

Barron, L. A. 2003. Ask and you shall receive? Gender differences in negotiators' beliefs about requests for a higher salary. *Human Relations* 56 (6): 635–662.

Bazerman, M. H., G. Loewenstein, and D. A. Moore. 2002. Why good accountants do bad audits. *Harvard Business Review* (November): 96–103.

Braun, K. 2001. The disposition of audit-detected misstatements: An examination of risk and reward factors and aggregation effects. *Contemporary Accounting Research* 18 (1): 71–99.

Brown, H. L., and K. Johnstone. 2005. The effects of engagement risk and auditor negotiation experience on the process and outcomes of client-auditor negotiation. Working paper, Boston College.

———, and A. Wright. 2008. Negotiation research in auditing. *Accounting Horizons* 22 (1): 91–109.

Chapman, G. B., and E. J. Johnson. 1999. Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes* 79 (2): 115–153.

Cialdini, R. B., and K. Ascani. 1976. Test of a concession procedure for inducing verbal, behavioral, and further compliance with a request to give blood. *The Journal of Applied Psychology* 61: 295–300.

——— 2001. *Influence*. 3rd edition. New York, NY: HarperCollins College Publishers.

Cotter, M. J., and J. A. Henley. 2008. First-offer disadvantage in zero-sum game negotiation outcomes. *Journal of Business-To-Business Marketing* 15 (1): 25–44.

Flynn, F. J., and D. R. Ames. 2006. What's good for the goose may not be as good for the gander: The benefits of self-monitoring for men and women in task groups and dyadic conflicts. *The Journal of Applied Psychology* 91 (2): 272–281.

Galinsky, A. D., and T. Mussweiler. 2001. First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology* 81 (4): 657–669.

Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* (December): 535–563.

———, S. McCracken, and S. Salterio. 2005. Negotiations over accounting issues: The congruency of audit partner and chief financial officer recalls. *Auditing: A Journal of Practice & Theory* 24 (Supplement): 171–194.

———, ———, and ———. 2007. The chief financial officer's perspective on auditor-client negotiations. *Contemporary Accounting Research* 24 (2): 387–422.

Gillan, K. 2007. Unconscious human nature affecting "professional skepticism." Keynote Address at the Forum on Auditing in the Small Business Environment. Santa Monica, CA. January 23. Available at: http://pcaobus.org/News/Speech/Pages/01232007_GillanAuditingSmallBusinessEnvironment.aspx.

Hatfield, R., C. Agoglia, and M. Sanchez. 2008. Client characteristics and the negotiation tactics of auditors: Implications for financial reporting. *Journal of Accounting Research* 46 (5): 1183–1208.

Joe, J., A. Wright, and S. Wright. 2008. The impact of changes in the reporting environment on the disposition of proposed audit adjustments. Working paper, Northeastern University.

Johnstone, K. M., J. C. Bedard, and S. F. Biggs. 2002. Aggressive client reporting: Factors affecting auditors' generation of financial reporting alternatives. *Auditing: A Journal of Practice & Theory* 21 (1): 47–65.

Kristensen, H., and T. Garling. 2000. Anchoring induced biases in consumer price negotiations. *Journal of Consumer Policy* 23 (4): 445–460.

Lambert, T. A., J. F. Brazel, and K. L. Jones. 2009. Unintended consequences of accelerated filings: Are mandatory reductions in audit delay associated with reductions in earnings quality? Working paper, University of Massachusetts.

Lord, C. G., M. R. Lepper, and E. Preston. 1984. Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology* 47 (6): 1231–1243.

McCracken, S., S. Salterio, and R. Schmidt. 2008. Effects of differential experience on auditors' intended usage of negotiation strategies: Methodological and practice concerns. Working paper, McMaster University.

Messier, W. F., Jr., S. M. Glover, and D. F. Prawitt. 2008. *Auditing and Assurance Services: A Systematic Approach*. Boston, MA: McGraw-Hill Irwin.

Mussweiler, T., F. Strack, and T. Pfeiffer. 2000. Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin* 26 (9): 1142–1150.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings-management decisions. *The Accounting Review* (Supplement): 175–202.

Ng, T., and H. Tan. 2003. Effects of authoritative guidance availability and audit committee effectiveness on auditors' judgments in an auditor-client negotiation context. *The Accounting Review* 78 (July): 801–818.

———, and ———. 2007. Effects of qualitative factor salience, expressed client concern, and qualitative materiality thresholds on auditors' audit adjustment decisions. *Contemporary Accounting Research* 24 (4): 1171–1192.

Pruitt, D. G., and J. L. Drews. 1969. The effect of time pressure, time elapsed, and the opponent's concession

rate on behavior in negotiation. *Journal of Experimental Social Psychology* 5: 43–60.

——— 1981. *Negotiation Behavior*. New York, NY: Academic Press.

Pruitt, D. G., and P. Carnevale. 1993. *Negotiation in Social Conflict*. Pacific Grove, CA: Books/Cole Publishing.

——— 1998. Social conflict. In *The Handbook of Social Psychology*, 2, edited by D. T. Gilbert, S. T. Fiske, and G. Lindzey, 470–503. Boston, MA: The McGraw-Hill Companies, Inc.

Puncel, L. 2008. *Audit Procedures*. Chicago, IL: CCH, Inc.

Rubin, J. Z., and B. R. Brown. 1975. *The Social Psychology of Bargaining and Negotiation*. New York, NY: Academic Press.

———, and M. R. DiMatteo. 1972. Factors affecting the magnitude of subjective utility parameters in a tacit bargaining game. *Journal of Experimental Social Psychology* 8 (5): 412–426.

Sanchez, M., C. Agoglia, and R. Hatfield. 2007. The effect of auditors' use of a reciprocity-based strategy on auditor-client negotiations. *The Accounting Review* 82 (1): 205–240.

Securities and Exchange Commission (SEC). 1999. *Materiality*. Staff Accounting Bulletin No. 99. Washington, D.C.: SEC.

Tan, H. 1995. Effects of expectations, prior involvement, and review awareness on memory for audit evidence and judgment. *Journal of Accounting Research* 33 (1): 113–135.

Tan, H., and K. T. Trotman. 2009. Effects of auditors' concession timing on financial officers' negotiation judgments. Working paper, Nanyang Technological University.

Trotman, K. T., A. Wright, and S. Wright. 2005. Auditor negotiations: An examination of the efficacy of intervention methods. *The Accounting Review* 80 (1): 349–367.

———, ———, and ———. 2009. An examination of the effects of auditor rank on pre-negotiation judgments. *Auditing: A Journal of Practice & Theory* 28 (1): 191–203

Wansink, B., R. J. Kent, and S. J. Hoch. 1998. An anchoring and adjustment model of purchase quantity decisions. *Journal of Marketing Research* 35 (1): 71–81.

Wright, A., and S. Wright. 1997. An examination of factors affecting the decision to waive audit adjustments. *Journal of Accounting, Auditing & Finance* 12 (Winter): 15–36.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1669–1691

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1669

Attracting Creativity: The Initial and Aggregate Effects of Contract Selection on Creativity-Weighted Productivity

Steven J. Kachelmeier
Michael G. Williamson
The University of Texas at Austin

ABSTRACT: Performance-based compensation contracts can affect productivity both by motivating effort and by attracting workers whose abilities align best with the offered contracts. Using an experiment in which participants design "rebus puzzles," this study tests whether the incremental benefits of contract self-selection extend to a multi-dimensional performance environment in which participants choose between a contract that rewards both quantity and creativity versus a contract that rewards quantity only. Findings indicate that participants who choose a creativity-weighted pay scheme perceive themselves to have greater creative potential than those who choose a quantity-only scheme. This perceived creativity advantage manifests itself in the superior *initial* creativity ratings of such participants' productive output. However, in the *aggregate*, participants paid only for quantity eventually surpass the creativity-weighted productivity scores of participants paid for creativity-weighted productivity, whether compensation contracts are self-selected or randomly assigned. Thus, the implications of contract selection on creativity-weighted productivity hinge on the importance of the initial advantage enjoyed by participants who self-select a creativity-weighted contract.

Keywords: *creativity; incentives; contingent compensation; multi-dimensional performance measurement; adverse selection.*

We received helpful suggestions on previous drafts from Marlys Lipe (editor), two anonymous reviewers, Sarah Bonner, Larry Brown, Steve Glover, Lynn Hannan, Jessen Hobson, Bob Libby, Bill Messier, Mark Nelson, Tatiana Sandino, Jen Winchel, participants at the 2008 American Accounting Association Annual Meeting, Management Accounting Section Midyear Meeting, the Boston Accounting Research Colloquium, and workshop participants at Brigham Young University, University of Connecticut, Cornell University, Emory University, Florida State University, Georgia State University, Harvard University, University of Notre Dame, University of Pennsylvania, University of Southern California, the Negotiation Research Center of The University of Texas at Dallas, Tilburg University, and Washington University in St. Louis. We also thank Morrie Schulman of The University of Texas at Austin Division of Instructional Innovation and Todd Pinney of Turning Technologies, Inc. for providing and tailoring the application of radio-frequency response devices for use by our creativity raters. We gratefully acknowledge funding from a McCombs School of Business Research Excellence Grant and support to the first author from the Charles T. Zlatkovich Centennial Professorship and from the Randal B. McDonald Chair in Accounting.

Editor's note: The original submission for this article occurred before the transition to the current editorial regime, and was handled throughout the revision process by former editor Marlys Lipe, who accepted the most recent revision for publication.

Submitted: February 2008
Accepted: November 2009
Published Online: August 2010

Data Availability: *Please contact the authors.*

I. INTRODUCTION

In addition to motivating effort toward desired output characteristics, performance-based compensation schemes can potentially *attract* desired characteristics by appealing most to individuals whose abilities align best with the terms of the contract (Demski and Feltham 1978; Sprinkle and Williamson 2007, §2). Although several studies have used the tools of experimentation to demonstrate contract-selection benefits of performance-based pay by allowing participants to choose among competing contracts in laboratory production exercises (e.g., Chow 1983; Waller and Chow 1985; Shields and Waller 1988; Dillard and Fisher 1990; Cadsby et al. 2007), those experiments have been restricted to a single performance dimension—typically the quantity produced. The objective of the current study is to investigate whether contract-selection benefits extend to a multi-dimensional performance environment in which the firm desires both *quantity* and *creativity*.

Our interest in the contract-selection effects of creativity-weighted compensation is motivated by the importance of creativity to success in contemporary business settings (Chang and Birkett 2004; Fallon and Senn 2006). An added benefit of our focus on creativity is that it enables us to build on recent related work by Kachelmeier, Reichert, and Williamson (2008); hereafter, KRW), who investigate the effort-influencing effects of creativity-based compensation in an environment in which contracts are randomly assigned. KRW reach the provocative conclusion that creativity-weighted pay actually lowers creativity-weighted productivity, insofar as experimental participants who are paid only for quantity produce as much high-creativity output as their creativity-weighted counterparts, while producing significantly more output overall. KRW's conclusion is limited, however, by the fact that their experimental participants are randomly assigned to contracts, thereby preventing the creativity-weighted contract from attracting those with more creative ability. To the extent that both effort and ability drive creativity, addressing this limitation could potentially reverse KRW's conclusion, making it important to consider the effects of endogenous contract selection.

To pursue this objective, we adapt the "rebus puzzle" task used in KRW, in which experimental participants design puzzles that use words, symbols, and pictures to represent a familiar term or phrase. Rebus puzzles serve as an experimental analog to real-world settings involving the "production of ideas," for which firms benefit from both the quantity of production and its creative value. Prototypical examples of such settings include professional service firms (Chang and Birkett 2004) and universities (Dewett and Denisi 2004), among others. Within such a setting, our experiment manipulates (1) whether participants are compensated based on the raw quantity of puzzles produced or on quantity weighted by creativity scores from a panel of creativity raters, and (2) whether these two compensation contracts are assigned randomly or are chosen by the participants.

Findings indicate that, relative to the random-assignment control condition, self-selection does indeed attract the most creative participants to a creativity-weighted contract, both in terms of self-perceived creative potential and in terms of the actual creativity of participants' initial production. However, notwithstanding the benefit of self-selected creativity-weighted compensation that we observe in *initial* creativity, results from *aggregate* production reveal that participants paid only for quantity eventually catch-up with and surpass the creativity-weighted scores of those paid for creativity-weighted productivity, whether contracts are self-selected or randomly assigned. Thus, creativity-weighted compensation does not maximize (eventual) creativity-weighted productivity. We interpret the aggregate results as reflecting the eventual net effect of two opposing forces. Namely, while self-selection allows those who choose a creativity-weighted contract to

enjoy a significant initial advantage, the dysfunctional effort-influencing effects of creativity-weighted compensation eventually dominate.

While it might seem that the eventual dominance of quantity-only compensation demonstrates the robustness of KRW's similar conclusions to contract selection, we submit that the initial benefit of self-selected creativity-weighted compensation is important in its own right. Our self-selected quantity-only participants had the opportunity to make up their initial deficit by outperforming their creativity-weighted counterparts in aggregate production, but real-world environments are often not as forgiving. That is, in firms and organizations, resources often flow to those who achieve initial success, such that initial advantages can govern long-term opportunities. For this reason, we believe that our results add a significant caveat to KRW's conclusions about the superiority of paying only for quantity even when both quantity and creativity are desired. Namely, when employees can choose the contractual opportunities that suit them best, those who prefer creativity-based contracts could be faster "out of the gate."

We describe the background of our study and develop hypotheses in Section II. Section III explains the method and design, Section IV presents results, and Section V concludes.

II. BACKGROUND, THEORY, AND HYPOTHESES

Background: Creativity-Weighted Productivity

Accountants' interest in performance-based pay schemes stems from the assumption that firms cannot directly observe employees' (1) effort or (2) abilities (type). Basing compensation on output measures mitigates both sources of information asymmetry, insofar as employees can work harder to attain rewarded measures, thereby mitigating the moral-hazard problem of hidden effort, and can also pursue employment contracts that align best with innate skills, thereby mitigating the adverse-selection problem of hidden abilities (Demski and Feltham 1978). Sprinkle and Williamson (2007, §2) review several management accounting experiments that demonstrate these benefits. However, the studies they review are generally limited to one performance dimension—typically the quantity produced. Among others, Kaplan and Norton (1996) argue that the full potential of performance-based measurement can be captured only by generalizing to multiple dimensions.

As a step toward this objective, KRW extend the experimental study of performance-based compensation to include creativity in addition to quantity. Their interest in creativity is motivated in part by an apparent paradox: despite the clear importance of creativity to contemporary business (e.g., Fallon and Senn 2006), Ittner et al. (1997) report that very few firms include explicit measures of creativity or innovation in actual performance-based compensation schemes. One plausible reason why firms are reluctant to reward creativity is that creativity is difficult to measure, but KRW assert a more fundamental reason: the effort-directing effects of creativity-based pay could do more harm than good. In an experimental setting that uses a panel of independent raters to ensure reliable measures of creativity, KRW find that creativity-weighted productivity (i.e., quantity weighted by creativity ratings) is significantly higher among participants compensated only for quantity than among participants compensated for creativity-weighted productivity.[1] By way of explanation, KRW draw on the cognitive difficulty people face when making trade-offs between conflicting incentives (Payne et al. 1993, Ch. 2). A common decision heuristic in such settings is to prioritize on one criterion as a condition that must be met first, thereby leading to suboptimal trade-offs. Extended to the trade-off between creativity and quantity, this reasoning suggests that creativity-weighted compensation leads participants to restrict production to ideas

[1] KRW also test a fixed-pay control condition and a condition that rewards average creativity only (i.e., no reward for quantity), finding that creativity-weighted productivity under these contracts is even lower.

that they feel will be highly creative. However, the literature on creativity in psychology and management (e.g., Amabile 1996) documents that creativity does not necessarily emerge simply from trying harder. Accordingly, while the creativity-focused effort induced by a creativity-weighted pay scheme significantly lowers overall production, it does not enhance creative output.

KRW's finding is provocative, but is restricted to the effort-influencing effects of creativity-weighted compensation. We draw on KRW as a baseline, replicating their use of randomly assigned contracts as a benchmark for considering the possibly countervailing effects of contract selection. Proceeding in this manner enables us to isolate the incremental effects of contract selection under *ceteris paribus* conditions.

Initial Creativity: Does Creativity-Weighted Compensation Attract the Most Creative Workers?

Random assignment to treatment conditions is the cornerstone of experimental design, allowing researchers to separate treatment effects from the potentially confounding influence of self-selection biases (Shadish et al. 2002). However, for the interest management accountants have in performance-based compensation, self-selection is more than a "confound." Of specific relevance to our setting, even if creativity is unresponsive to increased *effort*, the possibility of different creative *abilities* could serve as a beneficial consequence of creativity-weighted pay, if more creative individuals are more attracted than others to a contract that rewards creativity.

Waller and Chow (1985) model the self-selection benefits of performance-based compensation schemes as a function of prospective employees' perceived abilities and preferences. Consistent with this model, several experiments have demonstrated contract-selection benefits for different ways to reward quantity (e.g., Chow 1983; Waller and Chow 1985; Shields and Waller 1988; Dillard and Fisher 1990; Cadsby et al. 2007). We are unaware, however, of any extensions to a multi-dimensional environment. For contract-selection benefits to extend to creativity in particular, two assumptions must be met. First, individuals must be aware of meaningful variation in their creative potential. Second, more creative individuals must be more willing than others to choose a contract that rewards creativity explicitly.

Redmond et al. (1993) provide some evidence in support of the first assumption, insofar as they find that individuals have reasonable self-insight into their creativity. Their evidence is from a single creative task that may or may not generalize to our setting. Still, the available evidence (also see Amabile 1996) suggests that creative abilities and corresponding self-insights differ sufficiently to afford a potential for contract selection benefits to occur.

We are unaware of any direct evidence regarding the second assumption that more creative individuals would be more inclined to select a creativity-weighted contract. More generally, however, Vaughn (1999) provides experimental evidence that individuals employ a variant of the "availability heuristic" in forming self-assessments of likely success (also see Tversky and Kahneman 1973; Kahneman and Frederick 2002; Schwarz and Vaughn 2002). Vaughn's (1999) reasoning suggests that individuals contemplating the benefits of a creativity-weighted contract would likely reflect on how easily some creative ideas come to mind. If so, we should see some difference in initial creativity that favors the creativity-weighed contract, reflecting the initial ideas that lead individuals to choose such a contract. Moreover, if such an effect reflects self-selection on ability rather than contractually induced effort, we should not observe a similar difference under randomly assigned contracts. We test this reasoning as our first hypothesis:

H1: Relative to any difference observed under randomly assigned contracts, when individuals are allowed to choose between contracts, those choosing a creativity-weighted pay scheme will evidence greater initial creativity than will those choosing a scheme that rewards quantity only.

Aggregate Creativity-Weighted Productivity

Importantly, H1 predicts a contract-selection effect in *initial* creativity, when selection effects are most likely to be observable. Over time, any selection benefits of a creativity-weighted contract become intermixed with the effort-directing effects of such a contract. That is, even if individuals of different types are attracted to different contracts, those contracts subsequently motivate different kinds of effort that can exacerbate, offset, or even reverse the initial differences. As summarized earlier, KRW find evidence from a randomized experiment that a creativity-weighted compensation contract generates lower creativity-weighted productivity than a quantity-only scheme, defining "creativity-weighted productivity" as the sum of the creativity ratings assigned to all output produced.[2] They assert that this effect arises because the creativity-weighted contract leads individuals to fixate on creativity to the detriment of overall production (and without realizing gains in high-creativity production). Hence, the harmful effect of a creativity-weighted contract on aggregate production effort, as documented by KRW, works in the opposite direction of the initial selection benefit of a creativity-weighted contract that we predict in H1. *Ex ante*, it is unclear which effect would dominate over the course of aggregate production in a setting with self-selected contracts, motivating a two-tailed test of the following null hypothesis:[3]

H2 (null form): In aggregate production and with self-selected compensation contracts, creativity-weighted productivity will not differ between individuals who select a creativity-weighted pay scheme and individuals who select a quantity-only scheme.

Stating H2 in terms of "creativity-weighted productivity" implies that all output "counts," although output of higher creativity counts more. However, some real-world settings may value *only* high-creativity output.[4] Thus, to corroborate and generalize our findings, we also examine the effects of our experimental treatments on the production of high-creativity efforts only, defined (similarly to KRW) as output in the overall top quartile of creativity ratings.

III. METHOD

Participants and Experimental Task

In each of 10 sessions, 8–10 volunteers participated in a 75-minute experiment, yielding a total of 90 participants. Similar to KRW, we recruited student participants from undergraduate core business classes, reflecting a variety of business majors. We also adapted KRW's basic task, subject to some modifications that we describe shortly. The task involved designing "rebus puzzles," defined in the instructions as a "kind of riddle in which words and/or diagrams are used to represent a familiar term or phrase," with 11 examples given as illustrations. Also consistent with KRW, the instructions common to all conditions observed that "while we do not place any rules on the kinds of rebus puzzles you can submit, we value both the number of different puzzles you construct (i.e., quantity) and the creativity of those puzzles (i.e., puzzles that are original ideas, innovative, and clever)." This instruction is important because our interest is in the incre-

---

[2] Arithmetically, "creativity-weighted productivity" is equivalent to the total quantity produced multiplied by the average creativity rating of that quantity. Experimental participants in our study and in KRW were provided with the simpler definition based on the sum of the creativity scores for all output produced.

[3] Our statement of H2 in the null form does not imply that we are predicting the null. Rather, we are testing the null against a two-tailed alternative, given the *ex ante* reasoning that a difference could arise in either direction.

[4] For example, paralleling the two conditions examined in this study, part of the Hollywood writers' strike of late 2007 involved writers' objections to television studios' proposals to "only pay for jokes that are funny, [thus departing from] the longstanding tradition of paying comedy writers by the word" (Queenan 2007, A16). In that setting, a joke that is not funny (insufficiently creative) might indeed add very little (or even negative) value at the margin.

mental effects of performance-based compensation, not the more basic goal-setting effects from knowing what matters to the experimenter (Shalley 1991; Bonner and Sprinkle 2002).

Participants worked independently at separate tables, designing puzzles on 3 × 5 inch index cards by drawing the puzzle on one side and its solution on the other side. Participants placed each completed puzzle in an individual output box, subject to the understanding that once in the box, a puzzle could not be removed.[5] This rule maintained control over the sequence of submissions, which is important to our goal of separately measuring initial and aggregate creativity-weighted productivity.

Design

We employ a 2 × 2 between-subjects experimental design that manipulates two factors: (1) *contract type*, basing performance-contingent compensation on quantity only or on creativity-weighted productivity, and (2) *contract determination*, with contracts either randomly assigned or chosen by the participants.

Our first experimental factor, contract type, holds constant the dependence of both compensation contracts on quantity, but varies whether the contract also rewards (i.e., weights) creativity. We favor this approach over other possibilities, such as offering a contract that rewards average creativity ratings, for two reasons. First, the scope of our study is restricted to an environment in which creativity is important, but not at the expense of overall productivity (Chang and Birkett 2004; Dewett and Denisi 2004). As discussed by KRW, without quantity objectives, workers can improve *average* creativity by pursuing the "denominator strategy" of reducing overall production without increasing high-creativity production. Such a strategy would be dysfunctional from the firm's perspective unless less creative output is actually harmful at the margin, so we provide quantity incentives in all conditions. Second, from an experimental design perspective, we wanted to simplify the participants' choice by highlighting creativity as the only construct differing between the two contracts. Notwithstanding this rationale, we recognize that some firms may place little value on low-creativity output, such that we conduct additional analyses to examine the sensitivity of our results to measures of high-creativity output only. Our second experimental factor, contract determination, addresses our primary objective of examining self-selected contracts, while replicating KRW's use of randomly assigned contracts as a control condition.[6]

Participants in all 10 sessions first read the same information about *both* incentive contracts. The instructions specify that those using the quantity-only contract will receive compensation based on a rate per puzzle that results in $5.00 for the participant (among all participants using this contract) submitting the fewest puzzles and $45.00 for the participant submitting the most puzzles, thereby yielding an average compensation around $25.00. The alternative, creativity-weighted contract specifies compensation based on the "creativity-weighted total score of all rebus puzzles" constructed, determined by adding the 1-to-10 creativity ratings assigned to those puzzles by "a group of students from a different class." The instructions in this version observe that "each puzzle you submit helps your total score, but higher rated puzzles count more." Similar to the quantity-only payment scheme, the creativity-weighted scheme anchors total compensation at $5.00 and

[5] Participants were allowed to hold on to a puzzle as long as they wished before submitting it to the output box, but substantially all participants submitted puzzles to the box sequentially as they finished them, with the few instances of discarded output reflecting incomplete efforts. This behavior is reasonable, given that each completed puzzle strictly benefits compensation under either a quantity-only or a creativity-weighted productivity pay scheme.

[6] Our random assignment condition is not an exact replication of KRW. Unlike their study, we provide participants in all conditions with (1) a practice phase, and (2) an explicit description of both contracts prior to determining one of these contracts for payment (randomly or by participant choice). These modifications facilitate participants' ability to select among the available contracts, so we keep these features constant in our random assignment condition for *ceteris paribus* comparability.

$45.00 for the lowest and highest creativity-weighted total scores, respectively, yielding an average payment around $25.00.[7]

To operationalize the between-subjects manipulation of whether participants are randomly assigned to one of the contracts or are allowed to choose, we conduct five experimental sessions within each condition, in a counterbalanced order. In the random-assignment condition, participants draw a wooden ball that is marked "A," representing the quantity-only contract, or "B," representing the creativity-weighted contract. As noted in the instructions, an equal number of balls of each type ensures that the probability of being assigned to either contract is 50 percent. In the self-selection condition, each participant reads the same information about both contracts and writes in a checkmark next to the preferred contract when ready. While contract selection does not provide any assurance of similar sample sizes, our participants were nearly evenly split, with 24 (21) of the 45 self-selection participants choosing the quantity-only (creativity-weighted) contract.

Procedure

After describing the task and the statement about valuing both quantity and creativity, but *before* describing the compensation schemes, the instructions announced a five-minute practice period "to help you gauge your quantity and creativity in constructing rebus puzzles." The intent of the practice period is to familiarize participants with the task before choosing between contracts, so we provide the same opportunity in the random-assignment condition for *ceteris paribus* comparability.

After the five-minute practice period, the experimenters collected all rebus puzzles constructed during this period and distributed new instructions about the compensated phase of the experiment. These new instructions informed participants that they would now "design rebus puzzles for an additional 20 minutes for money." To avoid ambiguity about duplicating ideas from the practice phase, the instructions informed participants that "you may reuse any of your ideas from the previous five-minute phase if you choose to do so, but you must prepare a new card for each such idea." Participants then read descriptions of both compensation schemes and were either randomly assigned to one of these contracts or indicated the preferred contract, as described above. Next, participants completed a pre-experimental questionnaire to confirm their understanding of the pay schemes and to elicit various self-perceptions (described later).

The 20-minute compensated production phase of the experiment followed. Participants then completed a post-experimental questionnaire, followed by a risk-preference exercise adapted from Boylan and Sprinkle (2001). The risk-preference instrument required participants to make 15 choices between a certain payment of $5.00 or a lottery that paid $10 with probability p and $0 with probability 1 – p, with p decreasing in 5-percent increments from 85 percent to 15 percent. As a measure of risk preferences, we recorded the point at which each participant's choice shifted from preferring the lottery to preferring the certain $5 (participants were allowed to shift only once). Earlier shifts are recorded as lower scores, reflecting lower tolerance for risk. After participants completed the questionnaire, we played out the lottery for one of the 15 choices, selected at random. We then added either the lottery outcome or the certain $5.00 to each participant's compensation, depending on that participant's choice.

---

[7] Anchoring the low and high payments for both contracts at $5 and $45, respectively, allows us to vary the *basis* of compensation while holding constant the expected average *magnitude* of compensation. Similar to KRW, we achieve this equivalence by basing the low and high payments on the lowest and highest performers within each contract. However, these contracts are not "tournament" schemes as defined by Bonner et al. (2000), because with the sole exception of the highest-scoring participant, every other participant can earn greater compensation by increasing the compensated performance measure, irrespective of ordinal ranking.

Approximately two weeks after the experimental sessions, we paid participants privately in accordance with their task performance, corresponding compensation scheme, and any additional earnings from the risk-preference exercise. As announced in the instructions, the delay in payments was necessitated by the need to determine quantity counts and creativity ratings as the basis for compensation. Payments were determined exactly as described, without deception of any form.

Determining Creativity Ratings

Undergraduate student volunteers (n = 40) from a different class rated the creativity of the rebus puzzles.[8] In each of four three-hour rating sessions, 10 different raters evaluated a subset of the 2,178 puzzles produced by the experimental participants (324 puzzles during the practice phase and 1,854 during the compensated phase). Raters were paid $50 for their time.

To construct the subset of rebus puzzles evaluated by each cohort of raters, we first removed all practice-phase puzzles that were reused by the participants during the compensated phase of the experiment. This procedure ensured that participants would not be penalized for reusing an idea from the practice period, as was explicitly allowed in the instructions. For analysis purposes, we assign the same rating to each removed practice puzzle as the rating of the corresponding compensated puzzle that reused the idea. This process reduced the total number of evaluated rebus puzzles from 2,178 to 1,968.

For rating purposes, we randomized the order of the 1,968 puzzles. To reduce rater fatigue and make the task manageable, each of the four rating sessions processed one-fourth of the randomly ordered puzzles. We have no reason to suspect any systematic differences between rating sessions, and even if such differences existed, they would be unlikely to interact with the treatment factors distributed within each rating session. Nevertheless, to assess reliability across rating sessions and to give each session the same starting point for calibration, we began each rating session with the same 60 puzzles, in the same order. Correlations between the average rating within each rating session and the average across the other three rating sessions are significantly positive for the first 60 puzzles, ranging from 0.79 to 0.89. We therefore believe that the creativity ratings are reasonably reliable and consistent across sessions.[9]

Each rating session took place in a classroom, in which raters used radio-frequency response device keypads to assign a 1 (lowest) to 10 (highest) creativity rating to each puzzle, projected one at a time using a digital document camera. An assistant transcribed the solution to each puzzle on the bottom of each card beforehand, such that the puzzle solutions were projected along with the puzzles. The response-device software informed us when all ratings had been entered into a computer database, upon which one of the experimenters projected the next puzzle. To ensure independent ratings, we spaced raters far apart and asked them to refrain from comments or laughter. All raters first read the task instructions given to the experimental participants, including the same instructional examples. Raters were unaware, however, of the different compensation schemes or the assignment of participants to those schemes.[10]

---

[8] Our use of undergraduate raters differs from KRW's use of doctoral students. We felt that undergraduate peer ratings would be just as reliable as doctoral-student ratings. The use of undergraduates allowed us to recruit a larger number of raters, given the need to evaluate approximately 2,000 puzzles for creativity. We have no reason to believe that the undergraduate ratings from our study differ systematically from the doctoral-student ratings in KRW, but the more important point is that we hold the rater pool constant across both the random-assignment and self-selection conditions, such that if we replicate KRW's findings in the random-assignment condition, we gain comfort that our contract-selection results are not simply an artifact of different creativity raters.

[9] For tests of hypotheses, we obtain a creativity rating for the first 60 puzzles rated in all four sessions by randomly assigning 15 of these puzzles to each of the four sessions.

[10] Unobtrusive identification codes on the back of each puzzle were not projected to raters.

Within each rating session, correlations between each rater and the average of the other nine raters are significantly positive (ranging from 0.20 to 0.73), with the exception of one rater who exhibits an insignificant negative correlation ($r = -0.06$). To enhance reliability, we drop the lowest correlated rater within each rating session. Cronbach's Alpha is 0.79 under this approach, which exceeds typical reliability thresholds (Peterson 1994).

As illustrations, Figure 1 provides examples of six rebus puzzles from our experiment, with four examples of high-creativity ratings (Panels A through D), one example of a puzzle at the median of the creativity distribution (Panel E), and one immediately above the bottom quartile (Panel F).[11] For the entire distribution of puzzles, average ratings (nine raters per puzzle) range from 1.44 to 8.89, reflecting a wide variation in creativity.

IV. RESULTS

Necessary Conditions for Contract Selection Benefits

Before testing our hypotheses, we first explore several necessary conditions for any potential benefit of contract self-selection. First, participants must perceive meaningful differences in their potential to produce highly creative puzzles, the criterion that differs between contracts, and select accordingly. Second, we must observe meaningful variation in actual creative production, such that potential self-selection benefits are detectable. Third, participants' self-assessed creativity potential should correlate positively with their actual creative production. Below, we test each condition within our contract self-selection subsample. Table 1 provides descriptive statistics, including means, standard deviations, and ranges by condition for the dependent measures.

Self-Assessed Creative Potential and Basis of Selection

We elicited participants' self-assessments of potential relative to others to produce highly creative rebus puzzles in a questionnaire administered after the practice session but before the compensated phase of the experiment. Participants responded on a nine-point Likert scale, with the low and high endpoints labeled "much worse than others" and "much better than others," respectively. As Table 1 highlights, participants' responses varied widely, ranging over the entire nine-point scale.

As noted earlier, nearly half of the participants in the contract-selection condition (21 of 45) selected the creativity-weighted contract. Univariate comparisons of means reveal that, relative to those selecting the quantity-only contract, participants choosing the creativity-weighted contract perceive themselves to have significantly more potential to produce highly creative puzzles ($t = 5.26$; two-tailed $p < 0.01$). Moreover, participants rating their creativity potential above the midpoint on this scale (the median score) predominately select the creativity-weighted contract over the quantity contract (16 creativity-weighted versus 3 quantity-only), whereas participants at or below the midpoint predominately select the quantity-only contract (5 creativity-weighted versus 21 quantity-only). These relative frequencies are significantly different ($\chi^2 = 18.62$; $p < 0.01$).

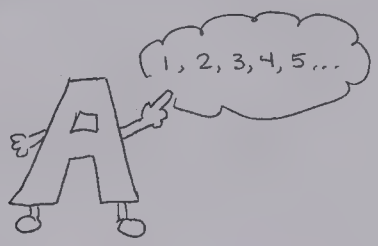
To corroborate our univariate results, we also perform a logistic regression analysis (untabulated) of the propensity to choose the creativity-weighted contract, regressed on self-reported creative potential, the supplemental measure from the post-experimental task we use to elicit risk preferences, and on the interaction between these two independent measures.[12] Controlling for risk

[11] The relatively low-rated puzzle in Panel F of Figure 1, "crossing paths," follows a pattern similar to one of the instructional examples, "cross roads." KRW observe that creativity raters tend to assign lower ratings to puzzles that are patterned after one of the instructional examples. Our experiment uses the same instructional examples as illustrated in KRW (2008, 369–370).

[12] We control for risk preferences because the subjectivity of creativity ratings implies greater uncertainty under the creativity-weighted pay scheme than under the quantity-only scheme. In general, individuals with greater tolerance for

FIGURE 1
Examples of Actual Rebus Puzzles

Panel A



Solution: Accounting
Creativity rating: 8.33
Condition: Random/ Quantity-only

Panel B



Solution: Carry a tune
Creativity rating: 7.89
Condition: Random/Creativity-weighted productivity

Panel C



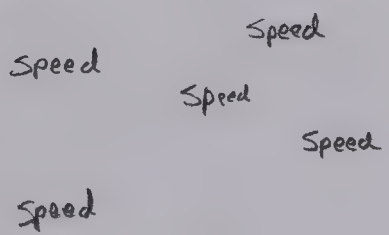
Solution: Nail biter
Creativity rating: 7.33
Condition: Selection/Quantity-only

Panel D



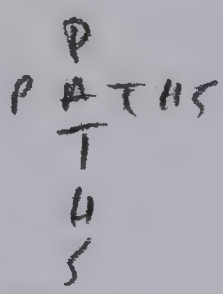
Solution: Spoiled milk
Creativity rating: 7.56
Condition: Selection/Creativity-weighted productivity

Panel E



Solution: Five speed
Creativity rating: 4.67 (Median Ranking)
Condition: Random/Quantity-only

Panel F



Solution: Crossing paths
Creativity rating: 3.78 (25th Percentile)
Condition: Selection/Quantity-only

risk are more likely to prefer noisier compensation schemes (Cadsby et al. 2007). Consistent with this reasoning, the propensity of participants to select the creativity-weighted contract is positively associated with risk preferences ($p = 0.05$). Also, the creative potential × risk preference interaction is significantly negative ($p = 0.05$), suggesting that the combined effect of the creative potential and risk-preference measures is subadditive: a greater score on either one of these measures dampens (but does not eliminate) the effect of the other measure on the propensity to select the creativity-weighted contract.

TABLE 1

Descriptive Statistics

	Self-Selected Contracts		Randomly Assigned Contracts	
	Mean (Standard Deviation) {Range}		Mean (Standard Deviation) {Range}	
	Quantity-Only	Creativity-Weighted	Quantity-Only	Creativity-Weighted
	n = 24	n = 21	n = 19	n = 26
Creativity Potential	3.92 (1.25) {2–7}	6.14 (1.59) {3–9}	5.26 (1.85) {3–9}	5.15 (1.89) {1–8}
Creativity Rating First Puzzle	4.61 (1.16) {2.11–7.22}	5.09 (1.34) {3.33–7.44}	4.80 (1.22) {2.78–6.78}	4.45 (1.17) {3.00–7.44}
Average Creativity Rating First Eight Puzzles	4.72 (0.59) {3.67–5.82}	5.09 (0.44) {4.42–5.89}	4.98 (0.62) {4.06–6.01}	4.87 (0.51) {3.54–5.83}
Number of High Creativity Puzzles through the Eighth Unit of Production	1.75 (1.36) {0–5}	2.52 (1.21) {0–5}	2.21 (1.62) {0–5}	2.08 (1.29) {0–5}
Aggregate Creativity-Weighted Productivity	113.32 (57.12) {34.56–254.22}	77.31 (29.92) {31.00–141.00}	122.05 (40.98) {57.56–186.78}	83.88 (27.07) {35.00–140.11}
Aggregate Number of High Creativity Puzzles	4.33 (2.84) {0–11}	4.48 (2.06) {1–10}	5.95 (4.01) {0–15}	4.69 (3.06) {1–12}
Quantity	25.13 (15.41) {7–76}	16.85 (6.06) {6–29}	25.74 (10.22) {12–54}	16.85 (5.42) {7–29}
Self-Perceived Task Familiarity	4.92 (2.60) {1–9}	4.86 (2.87) {1–9}	5.26 (2.81) {1–9}	5.19 (3.21) {1–9}

preferences, logit results confirm that the propensity of participants to select the creativity-weighted contract is positively associated with self-perceived creative potential ($p < 0.01$).[13]

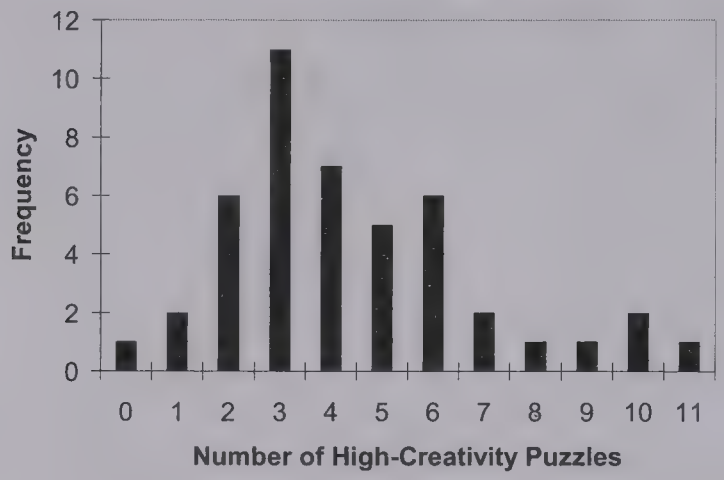
Meaningful Variation in High-Creativity Production

To assess whether our participants exhibit meaningful variation in creative production, we examine the production of high-creativity puzzles, defined similarly to KRW as the number of puzzles each participant produces that lie in the overall top quartile of creativity ratings, which in our distribution requires a rating of 5.7 or higher.[14] Figure 2 graphs the overall distribution of

[13] We observe no significant influence of the five-minute practice phase on contract preferences, and do not detect any significant differences in quantity, average creativity, or creativity-weighted productivity during practice. Participants likely needed the practice time to learn the task and experiment with different types of puzzles. Results are also insensitive to participant gender, with 55 percent (52 percent) of females (males) choosing the quantity-only contract.

[14] One might conjecture that average creativity ratings could serve as an alternate measure of creative production, but that measure is confounded by what KRW termed a "denominator effect." Namely, if two participants (hypothetically)

FIGURE 2
Distribution of High-Creativity Output by Participant in the Self-Selection Condition



high-creativity output for participants in the self-selection condition, collapsing across both possible contracts. As Figure 2 indicates, the variation of high-creativity production is quite dispersed, with several participants producing two or less such puzzles, while others produce ten or more.[15]

Correlation between Creative Potential and High-Creativity Output

Although participants choose contracts based on self-perceived creativity, and although we observe significant variation in high-creativity production, we *do not* observe a significant correlation between self-assessed creativity potential and the aggregate number of high-creativity puzzles produced ($r = 0.10$; $p = 0.51$). It appears that participants were unable to map their selection preferences to their aggregate capacity to produce high-creativity output.

Notwithstanding the absence of correlation between self-perceived creative potential and *aggregate* high-creativity production, we obtain a quite different result when we restrict the analysis to *initial* production, defined as the number of high-creativity puzzles generated within the first eight puzzles produced by each participant (eight is half the median quantity produced by participants in the self-selection condition). Here the correlation between self-perceived creative potential and initial high-creativity production is much larger, positive, and statistically significant ($r = 0.30$; $p = 0.05$).

By way of explanation, recall that the theory underlying H1 suggests that self-assessed creativity potential draws on initial ideas (Vaughn 1999). Hence, we observe a significantly positive correlation between self-perceived creativity and a measure of *initial* creative production, although

produce the same number of high-creativity puzzles, the participant who produces more puzzles overall will have a lower average creativity score, even though that participant is no less "creative" in an absolute sense, and in fact is more productive overall. We therefore prefer the raw count of high-creativity puzzles as a cleaner measure of creative production.

[15] Using the distribution in Figure 2, we also construct two "pseudo samples" by dividing the participants in our self-selection conditions into those who produced high-creativity output above or below the median (deleting participants at the median). The above-median group significantly exceeds the below-median group both in terms of (1) creativity-weighted productivity and (2) production of high-creativity puzzles ($p < 0.05$; two-tailed). This analysis confirms that we have enough statistical power to discern creativity-based advantages in aggregate high-creativity production, if participants in the self-selection condition have the foresight to select contracts that match their actual differences in creative production.

that correlation dissipates in *aggregate* creative production, likely reflecting the offsetting effort-inducing effects of the two contracts over time. This conclusion is consistent with the results of our hypothesis testing, as presented next.

Tests of Hypotheses

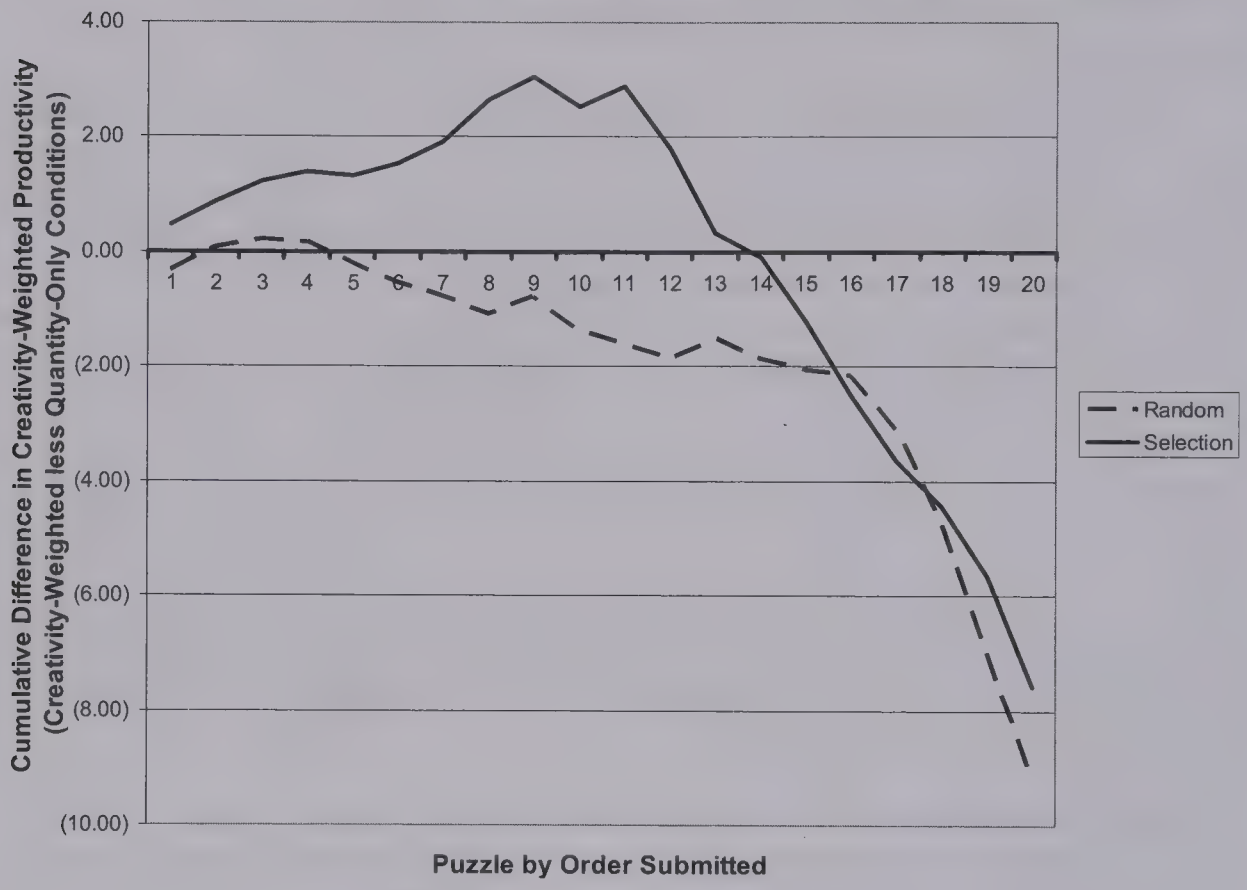
In the results presented below, we use analysis of covariance (ANCOVA) to investigate the main and interactive effects of contract type (quantity-only versus creativity-weighted productivity incentives) and contract determination (self-selection versus random assignment) on both initial creativity (H1) and aggregate creativity-weighted productivity (H2). The sole covariate in all analyses presented is a measure of participants' self-perceived familiarity with rebus puzzles, as captured from nine-point Likert scale responses to the post-experimental question, "I was familiar with rebus puzzles before the experiment," with the low and high endpoints labeled "strongly disagree" and "strongly agree," respectively. This covariate is statistically significant in all analyses reported below. Task familiarity is possible in this experiment from exposure to rebus puzzles in the popular media (e.g., Morris 1983). Extracting the variation from this covariate enhances our statistical ability to isolate creativity effects apart from any pre-experimental exposure to the task domain. However, task familiarity does not contaminate our ability to isolate contract effects, insofar as a separate ANOVA with task familiarity as the dependent variable (untabulated) reveals no significant main or interaction effects of either contract type or contract determination (highest $F = 0.31$; all $p > 0.50$). Hence, task familiarity appears evenly distributed across cells, including within the self-selection conditions.

H1: Initial Creativity

As explained in the motivation of H1 and consistent with the correlation analysis reported above on participants' self-perceived creativity, the selection benefits of a creativity-weighted contract are most likely to be manifest in *initial* creativity, before the possibly countervailing effects of differential effort can emerge across aggregate production (as examined later in H2). To illustrate the differences we observe across time, Figure 3 shows separate plots for the self-selection and random-assignment conditions of the *cumulative difference* in creativity-weighted productivity scores between those operating under the creativity-weighted contract and those paid for quantity only, at each unit of production. Positive (negative) values on the vertical axis indicate that the creativity-weighted contract generates higher (lower) cumulative creativity-weighted productivity than the quantity-only contract through the unit of production reflected on the horizontal axis. Thus, Figure 3 is essentially a running "horse race" of which contract is generating higher cumulative creativity-weighted productivity scores at each unit of production.

Consistent with H1, Figure 3 indicates that individuals selecting the creativity-weighted productivity contract enjoy an initial advantage, outperforming their quantity-only counterparts by a margin that increases in early production, before starting to reverse after the eighth unit of production. Statistically, the difference in favor of the creativity-weighted contract within the self-selection condition is significant for the first eight units ($t = 2.19$; one-tailed $p = 0.02$). The *interaction* between contract type and contract determination is also significant for cumulative creativity-weighted productivity scores through the eighth unit of production ($F = 4.82$; one-tailed $p = 0.02$), indicating that the difference favoring creativity-weighted pay within the self-selection condition for these early puzzles statistically exceeds the difference observed in the random-assignment condition.

FIGURE 3
Cumulative Creativity Weighted-Productivity by Order Submitted



For both the random assignment and self-selection conditions, this graph presents the difference in cumulative creativity-weighted productivity between individuals with creativity-weighted productivity incentives and those with quantity-only incentives by order of puzzle submitted.

Across the first eight puzzles, "creativity-weighted productivity" is essentially a scalar multiple of average creativity ratings, as nearly all of our participants produce at least eight puzzles.[16] For simplicity, we therefore shed further insight on H1 by examining average creativity ratings within initial production, noting that tests of initial creativity-weighted productivity for the same quantities would yield the same inferences. Starting with the creativity rating of the very first unit of production, Table 2 presents ANCOVA results (Panel B), simple effects (Panel C), and a graphical depiction of the means and the significant interaction detected in the ANCOVA (Panel A). Table 3 provides similar information for the average creativity ratings of participants' output across the first eight puzzles produced.

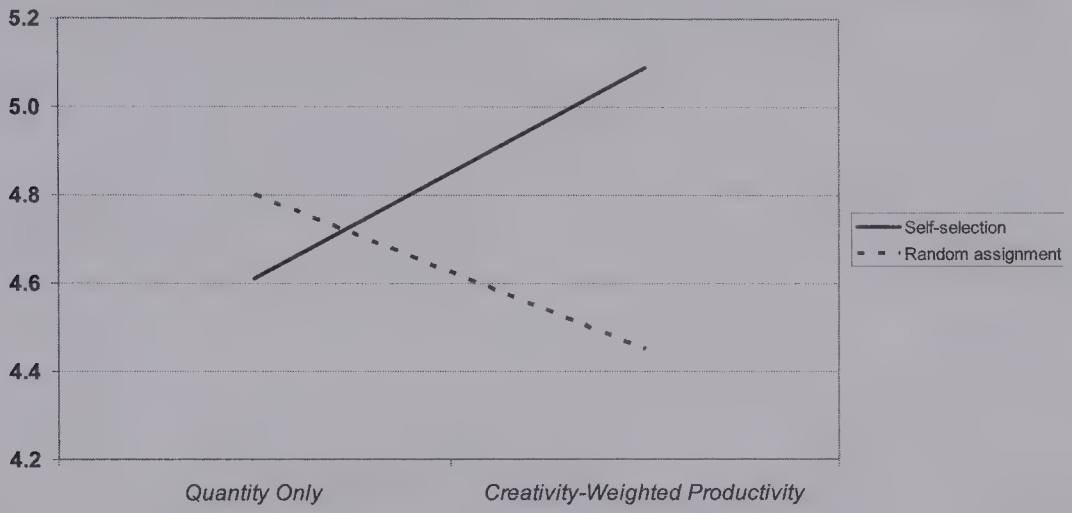
The first-puzzle results are subject to the noise inherent in creativity ratings for just one puzzle per participant, but even for the first puzzle (Table 2), we see a statistically significant interaction

[16] Only three of the 90 participants, each in a different experimental condition, produced less than eight puzzles, and no participant produced fewer than six puzzles.

TABLE 2

ANCOVA for the Effect of Contract Type and Contract Determination on the Creativity Rating of the First Puzzle Produced

Panel A: Mean Creativity Rating of First Puzzle



Panel B: Analysis of Covariance on Creativity Rating of First Puzzle

Factor	Df	Sum of Squares	F	p-value[a]
Contract Type	1	0.13	0.10	>0.50
Contract Determination	1	1.77	1.38	0.24
Contract Type × Determination	1	3.76	2.92	**0.05**
Task Familiarity (covariate)	1	18.81	14.62	<0.01
Error	85			

Panel C: Simple Effects

Factor	Df	Sum of Squares	F	p-value[a]
Effect of Contract Type within Self-Selection	1	1.24	2.09	**0.08**
Effect of Contract Type within Random Assignment	1	2.69	0.97	0.33

[a] Reported p-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.

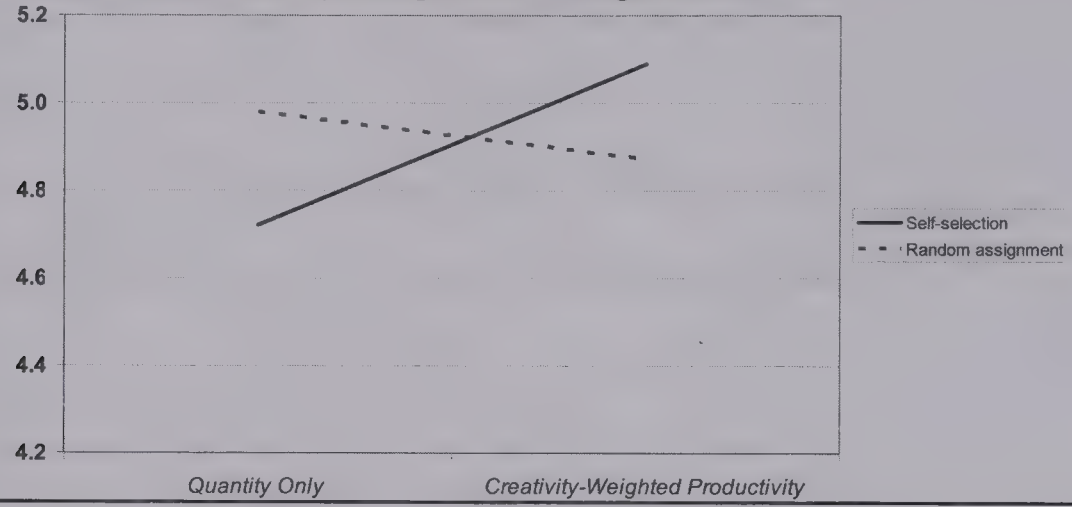
between contract type and contract determination ($F = 2.92$; $p = 0.045$) and at least marginally significant evidence that participants under the creativity-weighted contract exhibit higher creativity than those under the quantity-only contract when participants self-select the contract they prefer ($F = 2.09$; $p = 0.076$, one-tailed), but not when contracts are randomly assigned ($F = 0.97$; $p = 0.33$, two-tailed). We obtain stronger evidence for average creativity ratings across the first eight units of production, in which the interaction is highly significant ($F = 4.90$; $p = 0.02$) and the simple effects are clearly consistent with H1: average creativity ratings under the creativity-weighted contract are higher than those under the quantity-only contract only when participants self-select the contract they prefer ($F = 6.00$; $p < 0.01$, one-tailed), not when contracts are randomly assigned ($F = 0.49$; $p = 0.49$, two-tailed).

To be sure, defining "initial creativity" as average creativity ratings across the first eight units

TABLE 3

ANCOVA for the Effect of Contract Type and Contract Determination on Average Creativity Rating of the First Eight Puzzles

Panel A: Mean Average Creativity Rating of the First Eight Puzzles



Panel B: Analysis of Covariance on Average Creativity Rating of the First Eight Puzzles

Factor	df	Sum of Squares	F	p-value[a]
Contract Type	1	0.39	1.50	0.24
Contract Determination	1	<0.01	<0.01	>0.50
Contract Type × Determination	1	1.26	4.90	**0.02**
Task Familiarity (covariate)	1	3.22	12.50	<0.01
Error	85			

Panel C: Simple Effects

Factor	df	Sum of Squares	F	p-value[a]
Effect of Contract Type *within* Self-Selection	1	1.55	6.00	**<0.01**
Effect of Contract Type *within* Random Assignment	1	0.13	0.49	0.49

[a] Reported p-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.

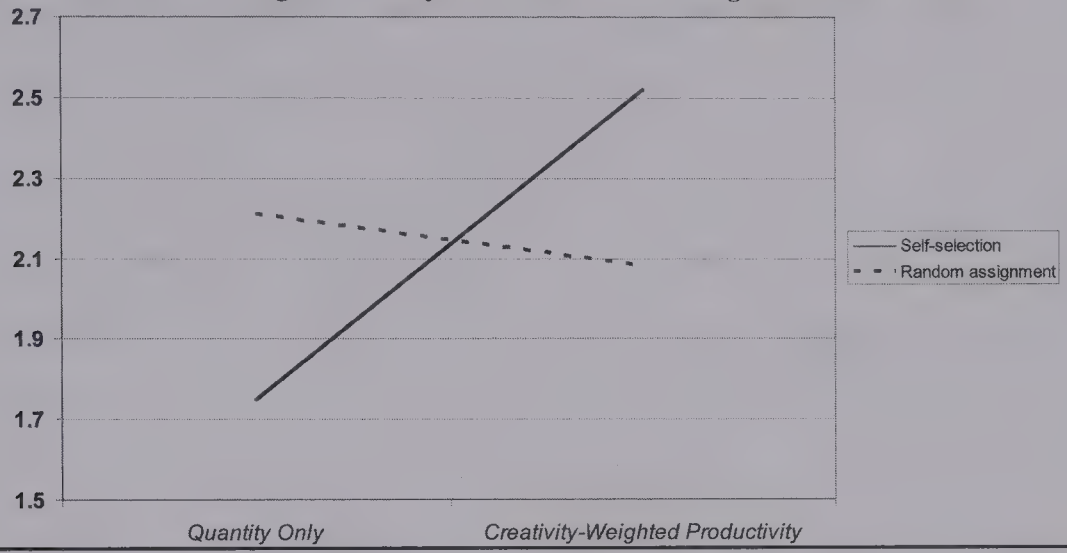
of production is somewhat *ad hoc*, but we obtain similar results defining initial creativity as the average creativity ratings of the first "n" puzzles, for all n between 1 and 8. That is, we find a statistically significant advantage in favor of the creativity-weighted contract when participants self-select (all F-statistics > 2.76; all $p \leq 0.05$), but not when contracts are randomly assigned (all $F < 0.42$; all $p > 0.50$). Overall, these results support H1; relative to the random assignment control condition, self-selection enabled those choosing the creativity-weighted contract to exhibit significantly higher initial creativity.

The above analysis reflects the creativity of all puzzles generated in early production. However, to the extent that many business settings place primary (or exclusive) value on high-creativity output, it is important to corroborate our findings with an analysis of high-creativity output only. Table 4 graphs means (Panel A), presents ANCOVA results (Panel B), and reports simple effects within the random-assignment and self-selection conditions (Panel C) for the num-

TABLE 4

ANCOVA for the Effect of Contract Type and Contract Determination on the Number of High-Creativity Puzzles in the First Eight Puzzles Produced

Panel A: Mean Number of High-Creativity Puzzles in the First Eight Puzzles Produced



Panel B: ANCOVA on Number of High-Creativity Puzzles in the First Eight Puzzles Produced

Factor	df	Sum of Squares	F	p-value[a]
Contract Type	1	2.44	1.51	0.22
Contract Determination	1	0.07	0.04	>0.50
Contract Type × Determination	1	4.54	2.81	**0.05**
Task Familiarity (covariate)	1	23.34	14.44	<0.01
Error	85			

Panel C: Simple Effects

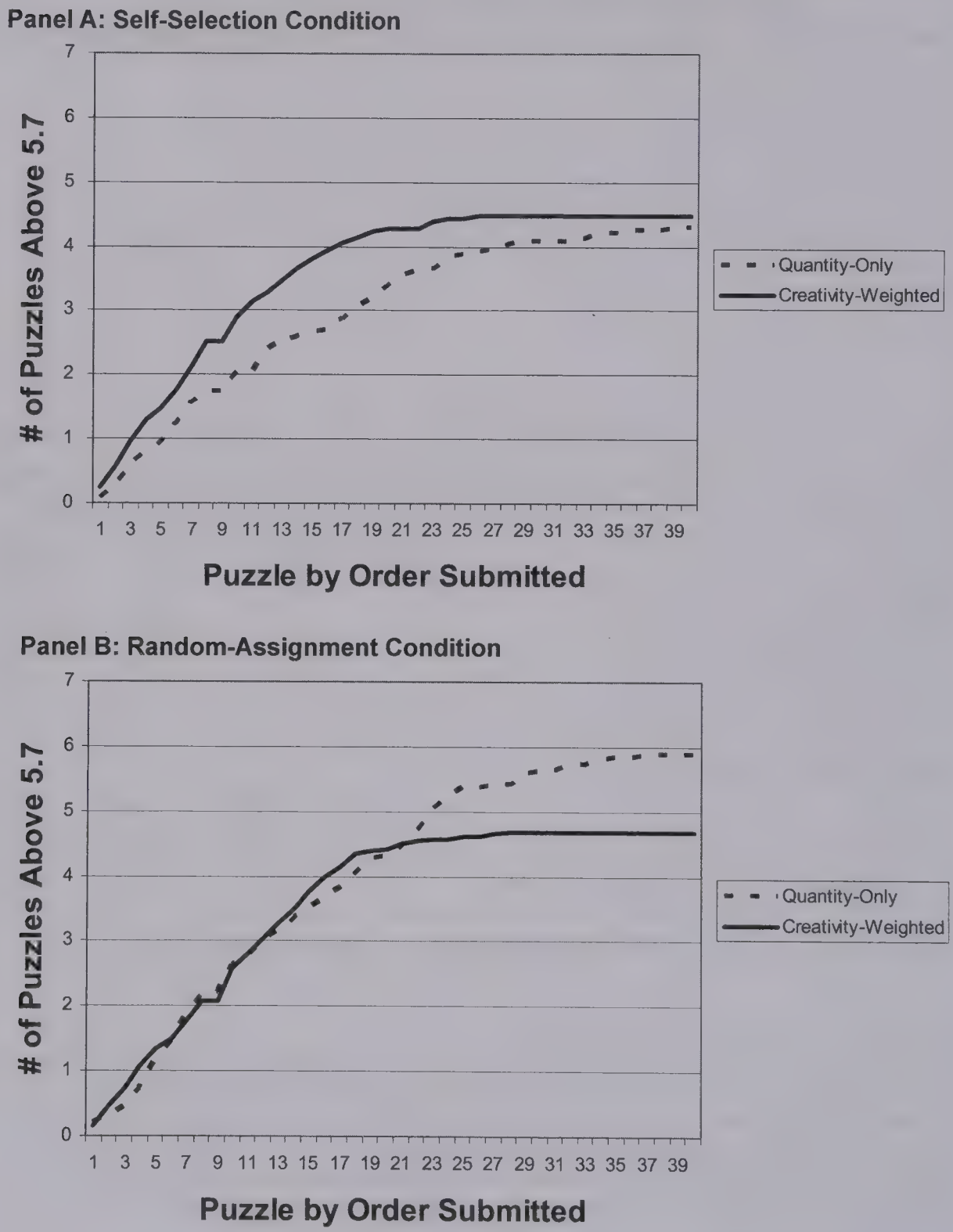
Factor	df	Sum of Squares	F	p-value[a]
Effect of Contract Type *within* Self-Selection	1	6.92	4.28	**0.02**
Effect of Contract Type *within* Random Assignment	1	0.17	0.10	>0.50

[a] Reported p-values are two-tailed unless testing a one-tailed prediction, as signified by bold face.

ber of high-creativity (i.e., overall top quartile) puzzles across the first eight units of production. As in Tables 2 and 3, the contract type × contract determination interaction in Table 4, Panel B is statistically significant ($F = 2.81$; $p = 0.05$). As Panel C indicates, participants using the creativity-weighted contract submitted more high-creativity puzzles through the eighth unit of production than quantity-only participants when participants self-select the contract they prefer ($F = 4.28$; $p = 0.02$, one-tailed), but not when contracts are randomly assigned ($F = 0.10$; $p > 0.50$, two-tailed).

To facilitate interpretation, Figure 4 depicts the pattern of high-creativity production across time. For self-selected contracts, those preferring the creativity-weighted scheme are "faster out of the gate," producing high-creativity puzzles at an initially higher rate than their quantity-only counterparts (Figure 4, Panel A). In the random-assignment condition, by contrast, the rate of

FIGURE 4
Cumulative Production of High-Creativity (Overall Top Quartile) Puzzles



initial high-creativity production is nearly identical between the creativity-weighted and quantity-only contracts (Figure 4, Panel B), suggesting that the temporal pattern in Panel A of Figure 4 is a contract-selection effect, consistent with H1.

H2: Aggregate Creativity-Weighted Productivity

Figures 3 and 4 illustrate that, despite their initial success, self-selected creativity-weighted participants eventually lose steam. That is, quantity-only participants eventually dominate in overall creativity-weighted productivity (Figure 3), and catch-up even in the number of high-creativity puzzles (Figure 4). To test H2, this section reports analyses for creativity-weighted productivity and high-creativity output aggregated over the entire production period.

Table 5 reports an ANCOVA with aggregate creativity-weighted productivity as the dependent variable.[17] The only significant effect from this ANCOVA is the main effect of contract type, with quantity-only participants significantly outperforming their creativity-weighted counterparts ($F = 19.06$; $p < 0.01$, two-tailed). The interaction between contract type and contract determination (self-selected versus random) is negligible ($F = 0.02$; $p > 0.50$). Thus, in aggregate creativity-weighted production, null hypothesis H2 must be rejected in favor of the quantity-only contract, irrespective of whether contracts are self-selected or randomly assigned.

To be sure, unlike the earlier tests of H1 in which production quantity is held constant, "creativity-weighted productivity" increases in both creativity *and* quantity across aggregate production. Accordingly, the superior creativity-weighted productivity scores of the quantity-only participants even in the self-selection condition could plausibly be consistent with such participants "flooding the market" with sufficient quantities of low-creativity output to offset their initial disadvantage in creativity ratings as detected earlier in H1. We address this possibility with two tests. First, we address the influence of low-creativity output on our results by repeating the aggregate creativity-weighted productivity ANCOVA after deleting all output in the overall *bottom* quartile of creativity ratings. Results (untabulated) remain similar to those reported above, with the main effect of contract type continuing to evidence greater creativity-weighted productivity among the quantity-only participants ($F = 12.85$; $p < 0.01$, two-tailed), with no main effect of contract

TABLE 5

ANCOVA for the Effect of Contract Type and Contract Determination on Creativity-Weighted Productivity

Factor	df	Sum of Squares	F	p-value (two-tailed)
Contract Type	1	30,174.76	19.06	<0.01
Contract Determination	1	961.48	0.61	0.44
Contract Type × Determination	1	24.81	0.02	>0.50
Task Familiarity (covariate)	1	6,920.83	4.37	0.04
Error	85			

[17] Because both creativity and quantity vary in aggregate production, we can no longer simplify the analysis by focusing on average creativity ratings as in testing H1. Average creativity ratings do not even serve as a reliable measure of "creativity" when quantity varies, due to the "denominator effect" that results in lower average creativity ratings with more overall production even if the production of high-creativity puzzles remains constant. Our subsequent tests in this section attempt to distinguish the effects of creativity and quantity on creativity-weighted productivity.

determination ($F = 1.09$; $p = 0.30$) and no interaction between contract type and determination ($F = 0.15$; $p > 0.50$).

A second approach is to analyze high-creativity production only—that is, the production of puzzles that fall in the overall top quartile of creativity ratings. ANCOVA results (untabulated) show that neither contract type, contract determination, nor the interaction between the two significantly influences the production of high-creativity puzzles in the aggregate (all $p > 0.20$).[18] Figure 5 illustrates this near equivalence, splitting the total quantity produced within each experimental condition into high-creativity puzzles and all other puzzles. This figure shows that quantity-only participants (both self-selected and randomly assigned) did indeed produce a much greater quantity overall, but did not do so at the expense of less high-creativity production. Thus, it would be an oversimplification to attribute the H2 results to quantity-only participants "flooding the market" with low-creativity efforts. As long as less creative output is not actually harmful at the margin, the quantity-only participants enjoy substantial gains in overall production while also producing just as many high-creativity efforts as those paid for creativity-weighted productivity.

FIGURE 5
Number of Puzzles with a Composite Creativity Rating Above and Below 5.7 by Experimental Condition



[18] Table 1 suggests that the high-creativity production in the quantity-only, random-assignment cell (5.95) could be higher than that of the other three cells (4,69, 4.33, and 4.48). However, none of the pairwise comparisons between individual cells (untabulated) are statistically significant at conventional levels (highest $t = 1.13$; two-tailed $p = 0.27$). Hence, the nominally higher production in the quantity-only, random-assignment cell could simply reflect chance variation.

As indicated earlier (see Figure 2), it is also unlikely that these results reflect a "ceiling effect" in high-creativity production, as that production exhibits considerable overall variance around the means depicted in Figure 5.

Returning to Figure 4, it becomes apparent how our H2 results reconcile to the earlier test showing that self-selected creativity-weighted participants produce significantly more high-creativity puzzles in initial production (H1). While these participants are indeed "faster out of the gate," Figure 4 shows that self-selected quantity-only participants continue to produce high-creativity puzzles over time, while their creativity-weighted counterparts appear to "burn out." We interpret this pattern as consistent with the opposing effects of two phenomena. Namely, the contract-selection treatment does appear to attract the most (initially) creative participants to the creativity-weighted contract, but over time, the effort-inducing effects of the two contracts lead even the participants who self-select the creativity-weighted contract to try too hard to be creative and too little to be productive.

V. CONCLUSIONS

In what we believe is the first experimental study to investigate contract selection in a multi-dimensional setting, we examine the effects on creativity-weighted productivity of allowing participants to choose between a contract that rewards creativity-weighted productivity and one that rewards quantity only. We find that participants who choose the creativity-weighted contract indicate higher self-perceptions of creative potential than do those who choose the quantity-only scheme. These self-perceptions are manifest in the initial creativity of participants' productive output, consistent with the premise that a creativity-weighted contract can attract relatively creative workers. To the extent that initial creative success can influence subsequent career paths and resources, this result qualifies KRW's conclusion from an experiment with randomly assigned contracts that creativity-weighted compensation impairs creativity-weighted productivity.

Despite the benefit of a self-selected creativity-weighted pay scheme that we observe in *initial* creativity, we find that quantity-only participants eventually surpass the creativity-weighted productivity scores of their creativity-weighted counterparts. Thus, in settings in which all workers have enough opportunity to "catch up," our results suggest that even those who self-select into a quantity-only contract can eventually produce just as many high-creativity efforts as their creativity-weighted counterparts, along with several other efforts of more modest creativity. Overall, the potential benefits of contract selection in a multi-dimensional setting that values both quantity and creativity appear to depend on which is most important: a fast start out of the gate (initial creativity) or sustainable production (aggregate creativity-weighted productivity). As we can envision environments of each variety, our results suggest a more complex answer to the costs and benefits of a creativity-weighted pay scheme than is provided in earlier research by KRW.

These results contribute to the contract-selection literature (e.g., Chow 1983; Waller and Chow 1985; Shields and Waller 1988; Dillard and Fisher 1990; Bouwens and van Lent 2006; Cadsby et al. 2007) in two important ways. First, our findings suggest a more subtle interpretation of the assumption implicit in many agency-based contract selection models that individuals have perfect insight into their abilities. For softer performance dimensions such as creativity, we find that individuals have reasonable self-insight into their *initial* creative abilities, but cannot necessarily anticipate the effort-inducing effects of the selected contracts in mapping this self-insight to sustainable aggregate performance. Thus, the descriptive validity of agency-based contract selection models is likely to depend on the nature of the ability organizations wish to attract. Second, our study highlights the challenges of achieving sustainable selection benefits in a multi-dimensional performance setting. Organizations must anticipate how contract design affects not only the attraction of high-skilled individuals, but also the (potentially suboptimal) trade-offs that the selected contracts motivate among multiple performance dimensions.

Our results also build on KRW's efforts to gain a better understanding of why firms are reluctant to reward creativity explicitly, despite calls to do so (Ittner et al. 1997). Certainly other elements besides contract selection could shed additional insights toward this end. For example, although we detect differences in initial versus aggregate production even in a 20-minute production exercise, truly creative efforts can often take days, weeks, or even years. A long-term experiment of this genre would be more difficult to control. Still, our results suggest that an internally valid longitudinal analysis of creativity-weighted productivity could shed important insights. Similar comments extend to the potential moderating effects of feedback (e.g., see Sprinkle 2008), albeit limited by the inherent delays that characterize reliable creativity feedback both in the laboratory and in practice. Finally, the results of a laboratory experiment with students producing "rebus puzzles" do not necessarily generalize to real-world environments with experienced professionals (such as marketing personnel) undertaking more elaborate creative tasks (such as designing advertising campaigns). Advertising, for example, might require a more unique approach to different client settings than is the case for designing different rebus puzzles. While it is not clear how any such differences would interact with our treatment conditions, theory-driven investigations of the effects of expertise or task characteristics on the efficacy of creativity incentives in a multi-dimensional performance setting could reveal important insights. We leave further exploration of these issues to future studies.

In closing, we echo KRW's call for expanding the experimental accounting literature on performance-based incentives beyond the traditional interest in quantity. Like KRW, our focus is on blending productivity with creativity, which we believe to be of particular importance in contemporary business practice. As a different example, Farrell et al. (2008) design an experiment to examine the quantity-quality trade-off, in which "quality" captures elements of accurate alignment with customer specifications. Whether creativity, quality, or other dimensions, we believe that accounting researchers can and should incorporate "softer" elements of performance into the study of performance-based compensation. Our study evidences that laboratory methods are well-suited to such inquiries for both the moral-hazard effects of hidden actions and the adverse-selection effects of hidden information.

REFERENCES

Amabile, T. M. 1996. *Creativity in Context: Update to the Social Psychology of Creativity*. Boulder, CO: Westview Press.

Bonner, S. E., R. Hastie, G. B. Sprinkle, and S. M. Young. 2000. A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Management Accounting Research* 12: 19–64.

———, and G. B. Sprinkle. 2002. The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society* 27 (May-July): 303–345.

Bouwens, J., and L. van Lent. 2006. Performance measure properties and the effect of incentive contracts. *Journal of Management Accounting Research* 18: 55–75.

Boylan, S. J., and G. B. Sprinkle. 2001. Experimental evidence on the relation between tax rates and compliance: The effect of earned vs. endowed income. *The Journal of the American Taxation Association* 23 (Spring): 75–90.

Cadsby, B. C., F. Song, and F. Tapon. 2007. Sorting and incentive effects of pay for performance: An experimental investigation. *Academy of Management Journal* 50 (April): 387–405.

Chang, L., and B. Birkett. 2004. Managing intellectual capital in a professional service firm: Exploring the creativity-productivity paradox. *Management Accounting Research* 15 (March): 7–31.

Chow, C. W. 1983. The effects of job standard tightness and compensation scheme on performance: An exploration of linkages. *The Accounting Review* 58 (October): 667–685.

Demski, J. S., and G. A. Feltham. 1978. Economic incentives in budgetary control systems. *The Accounting Review* 53 (April): 336–359.

Dewett, T., and A. S. Denisi. 2004. Exploring scholarly reputation: It's more than just productivity. *Scientometrics* 60 (June): 249–272.

Dillard, J. F., and J. G. Fisher. 1990. Compensation schemes, skill level and task performance: An experimental examination. *Decision Sciences* 21 (Winter): 121–137.

Fallon, P., and F. Senn. 2006. *Juicing the Orange: How to Turn Creativity into a Powerful Business Advantage*. Boston, MA: Harvard Business School Press.

Farrell, A. M., K. Kadous, and K. L. Towry. 2008. Contracting on contemporaneous vs. forward-looking measures: An experimental investigation. *Contemporary Accounting Research* 25 (Fall): 773–802.

Ittner, C. D., D. F. Larcker, and M. V. Rajan. 1997. The choice of performance measures in annual bonus contracts. *The Accounting Review* 72 (April): 231–255.

Kachelmeier, S. J., B. E. Reichert, and M. G. Williamson. 2008. Measuring and motivating quantity, creativity, or both. *Journal of Accounting Research* 46 (May): 341–373.

Kahneman, D., and S. Frederick. 2002. Representativeness revisited: Attribute substitution in intuitive judgment. In *Heuristics and Biases: The Psychology of Intuitive Judgment*, edited by Gilovich, T., D. Giffin, and D. Kahneman., 49–81. New York, NY: Cambridge University Press.

Kaplan, R., and D. Norton. 1996. *The Balanced Scorecard: Translating Strategy into Action*. Boston, MA: Harvard Business School Press.

Morris, S. 1983. *Omni Games: The Best Brainteasers from Omni Magazine*. New York, NY: Holt, Rinehart & Winston.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1993. *The Adaptive Decision Maker*. New York, NY: Cambridge University Press.

Peterson, R. A. 1994. A meta-analysis of Cronbach's coefficient alpha. *The Journal of Consumer Research* 21 (September): 381–391.

Queenan, J. 2007. No laughing matter. *Wall Street Journal* (November 12): A16.

Redmond, M. R., M. D. Mumford, and R. Teach. 1993. Putting creativity to work: Effects of leader behavior on subordinate creativity. *Organizational Behavior and Human Decision Processes* 55 (June): 120–151.

Schwarz, N., and L. A. Vaughn. 2002. The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. In *Heuristics and Biases: The Psychology of Intuitive Judgment*, edited by Gilovich, T., D. Giffin, and D. Kahneman., 103–119. New York, NY: Cambridge University Press.

Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin.

Shalley, C. E. 1991. Effects of productivity goals, creativity goals, and personal direction on individual creativity. *The Journal of Applied Psychology* 76 (April): 179–185.

Shields, M. D., and W. S. Waller. 1988. A behavioral study of accounting variables in performance-incentive contracts. *Accounting, Organizations and Society* 13 (November): 581–594.

Sprinkle, G. B., and M. G. Williamson. 2007. Experimental research in managerial accounting. In *Handbook of Management Accounting Research*, edited by Chapman, C. S., A. G. Hopwood, and M. D. Shields., Vol. 1. 415–444. Oxford, UK: Elsevier, Ltd.

——— 2008. Discussion of Measuring and motivating quantity, creativity, or both. *Journal of Accounting Research* 46 (May): 375–382.

Tversky, A., and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5: 207–232.

Vaughn, L. A. 1999. Effects of uncertainty on use of the availability heuristic for self-efficacy judgments. *European Journal of Social Psychology* 29 (March): 407–410.

Waller, W. S., and C. W. Chow. 1985. The self-selection and effort effects of standard-based employment contracts: A framework and some empirical evidence. *The Accounting Review* 60 (July): 458–476.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1693–1720

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1693

Seeking Shelter: Empirically Modeling Tax Shelters Using Financial Statement Information

Petro Lisowsky
University of Illinois at Urbana–Champaign

ABSTRACT: Using confidential tax shelter and tax return data obtained from the Internal Revenue Service, this study develops and validates an expanded model for inferring the likelihood that a firm engages in a tax shelter. Results show that tax shelter likelihood is positively related to subsidiaries located in tax havens, foreign-source income, inconsistent book-tax treatment, litigation losses, use of promoters, profitability, and size, and negatively related to leverage. Supplemental tests show that total book-tax differences (BTDs) and the contingent tax liability reserve are significantly related to tax shelter usage, while discretionary permanent BTDs and long-run cash effective tax rates are not. Finally, the model is weaker, yet still significant, in the FIN 48 disclosure environment. This research provides investors and policymakers with an extended, validated measure to calculate the presence of extreme cases of corporate tax aggressiveness. Such information could also aid analysts and other tax and non-tax researchers in assessing the benefits and risks of firm behavior.

Keywords: *tax shelter; FIN 48; book-tax differences; tax cushion; ROC Curve.*

Data Availability: *Financial data used in this study are available from public sources. Tax return data are obtained from the Internal Revenue Service's Large and Mid-Size Business Division (LMSB) Research & Workload Identification Office (RWI), and the Office of Tax Shelter Analysis (OTSA), and are not publicly available.*

I thank my dissertation committee at Boston University, Krish Menon (chair), Nishi Sinha, and Don Smith, for their skillful guidance in this research endeavor, as well as Paul Beck, Charles Boynton, Devan Mescall, Tom Omer, Leslie Robinson, and two anonymous referees for valuable input. I thank the participants at the Boston University Accounting Seminar Series, the 2008 University of North Carolina Tax Symposium, including Bob Adams of IRS (discussant), the University of Illinois at Urbana–Champaign, Stanford University, George Mason University, Lehigh University, University of New Hampshire, and University of Massachusetts Boston for helpful comments. I also thank the Internal Revenue Service for supporting this research by providing access to corporate tax return and tax shelter data. This study won the 2009 American Taxation Association/PricewaterhouseCoopers Outstanding Tax Dissertation Award.

Because tax data are confidential and protected by data nondisclosure agreements under the Internal Revenue Code, all statistics are presented in the aggregate. Any opinions are those of the author and do not necessarily reflect the views of the Internal Revenue Service.

Editor's note: Accepted by Thomas Omer.

Submitted: August 2008
Accepted: February 2010
Published Online: August 2010

I. INTRODUCTION

Tax shelters are transactions or arrangements that generate tax losses without incurring economic losses or risk (U.S. Department of the Treasury 1999, 6; hereafter Treasury). Their main goal is to lower the corporate tax liability by exploiting discontinuities in the tax law.[1] Tax shelters are illegal when they do not exhibit "economic substance" or a "business purpose"—that is, when they are created for the sole aim of evading tax rather than filling a non- (or pre-) tax economic need. From a research and tax administration standpoint, tax shelters are transactions that are classified as "Listed" or Non-Listed "Reportable" transactions by the U.S. Treasury Department (see Graham and Tucker 2006; Hennig et al. 2008). These transactions are described in specific Internal Revenue Service (IRS) notices, rulings, and guidance.[2] Inferring tax shelter use from financial statements is the main focus of this study.

The impact of tax shelters is substantial. The Treasury (1999) concluded that corporate tax shelters cost the federal government approximately $10 billion annually. In response to such tax base erosion, the IRS in 2000 established the Office of Tax Shelter Analysis (OTSA), whose sole task is to detect and track tax sheltering activity. Recent tax shelter litigation against KPMG challenged $2.5 billion in tax benefits (or $11 billion in sham tax losses) marketed and sold to clients from 1997 to 2001 (IRS 2005).

In light of these economic considerations, this study's main focus is to extend prior empirical work on inferring tax shelter participation from financial statement information (e.g., Wilson 2009). I use a conceptual framework from the Treasury (1999) white paper on tax shelters to extend the number and types of variables that can be used to infer corporate tax shelter use. Unlike previous studies, I am able to employ a relatively large, confidential tax shelter data set obtained from the IRS, allowing for the examination of 64 categories of tax shelters, with 267 tax shelter observations reported in the relatively short period from 2000 to 2004.[3] This analysis allows researchers, investors, and administrators to utilize clues in the financial statements to identify corporate tax shelter transactions.

In supplemental analyses, I examine the extent to which tax sheltering behavior is related to four alternative proxies of tax avoidance and tax aggressiveness that have recently emerged in the archival tax literature. In particular, I test whether tax shelters are related to discretionary permanent book-tax differences (Frank et al. 2009), long-run cash effective tax rates (Dyreng et al. 2008), total book-tax differences (Mills 1998), and the contingent tax liability reserve, or tax cushion (Gleason and Mills 2002). I also test the reported model's usefulness in gleaning tax sheltering behavior during the new regime of Financial Interpretation No. 48 (FIN 48, FASB 2006), in which firms must disclose unrecognized tax benefits. I conduct this analysis because it is unclear whether the new FIN 48 disclosures deter the use of corporate tax shelters, thus potentially reducing the power of the model developed here.

The ability to infer the presence of tax shelters is important for several reasons. Beyond lost tax revenues to the Treasury and related issues of equity in the tax system, identifying tax shelter firms can provide useful information about managerial risk-taking behavior (see Desai and Dharmapala 2006). The results of this study can also aid investors, analysts, and other researchers

---

[1] Graetz (1999, A1) more colorfully defines a tax shelter as a "deal done by very smart people that, absent tax considerations, would be very stupid."

[2] While the term "tax shelters" is used in the media to capture numerous types of general tax-saving strategies more commonly referred to in academic literature as "tax shields," such as depreciation and deferred compensation, the IRS's use of the term is more specific. See Sections II and III for details.

[3] This study does not claim to have access to the entire population of public U.S. firms engaging in tax shelters—only the firms whose tax shelters have been identified by IRS through audits or "voluntary" disclosures. There are obviously some tax shelters whose existence is yet to be discovered by the IRS.

interested in inferring tax shelter activities with respect to valuation (e.g., Hanlon and Slemrod 2009). Given greater regulatory scrutiny and significant penalties that could be imposed on firms, as well as new required financial disclosures for uncertain tax positions, knowledge of tax shelter use is increasingly relevant for investors and analysts as they assess the risk-reward trade-offs of investments; for auditors as they grapple with client reporting risks; for policymakers as they assess the current disclosure and compliance environment; and for tax administrators as they seek to supplement current compliance risk-assessment models. Furthermore, analyzing the extent to which empirical proxies for tax avoidance and aggressiveness represent underlying tax shelter transactions answers the call in Hanlon and Heitzman (2009) for researchers to develop a better understanding on the usefulness of such measures in various contexts.

The main results show that the likelihood of corporate tax shelter utilization is positively related to the presence of corporate subsidiaries located in tax havens, foreign-source income, inconsistent book-tax treatment, litigation losses, use of promoters,[4] profitability, and size. Tax shelter utilization is also negatively related to leverage. The model developed in this study increases the explanatory power of the Wilson (2009) model and corroborates the result in Graham and Tucker (2006) of the negative relation between debt and tax shelters. I also develop a metric called *TaxShelterScore*, which is derived from this study's main model parameter estimates. *TaxShelterScore* represents the estimated probability that a firm engages in a tax shelter, and can be used to infer corporate tax shelter usage. However, it cannot identify the dollar size of tax shelter benefits. Nevertheless, tests using *TaxShelterScore* show that the parameters have strong discriminatory power in inferring corporate tax shelter use. Applying *TaxShelterScore* to a holdout sample of firms, the area under the Receiver Operator Characteristic (ROC) Curve is 0.88. In other words, the model yields an 88 percent chance that a firm with a tax shelter has a higher predicted probability (or *TaxShelterScore*) than a firm without a tax shelter, for random pairs of firms with and without shelters. This result supports the model's use as a reliable measure for identifying the most severe form of tax aggressiveness—tax shelters.

To probe the relative benefits of the model's refinements in measuring tax shelter behavior in particular, supplemental analyses test the univariate link between tax shelters and four other proxies for tax avoidance and aggressiveness. Results show a strong positive relation between tax shelter usage, and both total book-tax differences and the contingent tax liability reserve, or tax cushion, while there is no significant association between tax shelter usage and either discretionary permanent book-tax differences or long-run cash effective tax rates. Meanwhile, the *TaxShelterScore* developed herein provides the strongest link to actual tax shelter incidence.[5] Finally, supplemental tests show that *TaxShelterScore* is weaker in identifying tax sheltering in the FIN 48 disclosure environment than in a pre-FIN 48 period.

The study is organized as follows. Section II presents a brief background on tax shelters, tax aggressiveness proxies, and FIN 48. Section III explains the OTSA tax shelter data and develops proxies for tax shelter characteristics. Section IV explains the sample and specifies the empirical model. Section V reports main results for the tax shelter likelihood model. Section VI reports supplemental analyses on the tax aggressiveness proxies and usefulness of the model during the FIN 48 disclosure period. Section VII concludes.

---

[4] The term "promoter" is used to represent the suppliers, marketers, and financiers of tax shelters.

[5] These results speak only to the various proxies' link to actual tax shelter incidence, but does not purport that the other measures do not capture more general (non-tax shelter) tax avoidance or tax aggressiveness.

II. BACKGROUND

Inferring Tax Shelters

The primary problem with tax shelters from research, investment, tax administration, and enforcement perspectives is detecting them. Because publicly traded corporate taxpayers disclose only their financial statements and not tax returns, it is difficult to calculate a firm's actual taxable income or tax burden with any degree of accuracy (Hanlon 2003), let alone identify the possible use of tax shelters.[6] McGill and Outslay (2004) detail the limitations of using only publicly available financial statements in detecting tax shelters. They analyze Enron's financial statements in an attempt to glean tax shelter activity, but acknowledge that it is extremely difficult to discern tax shelters by examining financial statements alone.[7]

Earlier research has examined patterns in financial statement information to infer tax shelter usage. For example, Plesko (2000) asserts that tax shelters are likely one reason why, in aggregate, pretax book income has grown at a faster rate than current or deferred tax expense from 1994 to 1998. Desai (2003) finds that, in 1998, less than half of book-tax differences are attributable to their main determinants of international operations, stock options, and depreciation accounting. Using a deductive approach, he suggests that the unexplained portion of book-tax differences is an indicator of general tax sheltering activity. Similarly, Desai and Dharmapala (2006) use a deductive approach to model management incentives to engage in tax shelters. This study proposes an inductive approach in identifying tax sheltering behavior.

More recent research identifies specific tax shelters using public sources, such as Tax Court cases and media news releases. Graham and Tucker (2006) use a sample of 43 firms identified from public records as having engaged in tax shelters over a 26-year period (1975–2000). They find that firms engaging in tax shelters use less debt, suggesting a substitution effect in which both vehicles can be used to lower taxable income. They acknowledge that their sample, which is based on public records, likely suffers from self-selection issues, and focuses on large cases that might not generalize to a broader population.

Wilson (2009) extends Graham and Tucker's (2006) sample to 2002 and develops a profile of firms that are likely to be engaged in tax shelters. He identifies 59 such firms over a 28-year period, with a total of 215 tax shelter firm-year observations. Wilson (2009) finds evidence that tax shelter participation is associated with book-tax differences, performance-matched discretionary accruals, size, return on assets (profitability), and foreign income.

My research directly extends Wilson (2009) by developing an expanded profile of tax shelter participants. I supplement his model with variables developed from a conceptual framework in the Treasury (1999) white paper on tax shelters. In designing the main tests, I use 267 tax shelter-year observations, as identified by the IRS's Office of Tax Shelter Analysis, over a relatively short five-year period of 2000–2004.[8]

---

[6] Prior literature also explains the limitations of using financial statements to estimate a corporation's general tax status (e.g., Hanlon and Shevlin 2002; McGill and Outslay 2002; Mills and Plesko 2003). Using IRS tax return data, Lisowsky (2009) develops an empirical model for inferring a corporation's tax burden from financial data.

[7] Tax return information that is not publicly disclosed, but that can help in identifying tax shelter usage includes information on Form 8886 and the Schedule M-3, which are used later in supplemental tests that encompass a more recent time period. However, identifying tax shelter transactions from tax returns alone can still remain elusive if the taxpayer is generally noncompliant in its tax filings to begin with.

[8] One advantage of using the OTSA data relative to the Graham and Tucker (2006) sample is that entry into the OTSA sample is not dependent on observations being discovered through tax litigation or public documents. One disadvantage of using tax return data is that differences in financial and tax consolidation rules can include or exclude certain entities or transactions for one set of books but not for another (Lisowsky and Trautman 2007; and Boynton et al. 2008). Therefore, as in Wilson (2009), I also include a broad measure for book-tax differences.

Alternative Measures of Tax Aggressiveness

A demand currently exists for a broadly accepted definition and measure of corporate tax aggressiveness (see Hanlon and Heitzman 2009). The difficulty is that tax aggressiveness is a matter of judgment, degree, and scope. In terms of judgment, firms, tax administrators, and even courts, must develop parameters around which transactions are deemed to exhibit economic substance or serve a business purpose. In terms of degree and scope, tax minimization can be conceptualized as a continuum, from avoidance to aggressiveness to sheltering.

I propose a broad definition of tax aggressiveness as behavior in which tax burden reduction is of first-order importance, with pre- or non-tax effects viewed as second-order, or marginal. Tax shelters, then, would be the most severe form of tax aggressiveness, as those transactions are clearly prohibited by the IRS. To illustrate, tax benefits from items such as depreciation or stock options are, in expectation, of second-order (or marginal) importance to the first-order capital allocation and managerial incentive alignment effects, respectively. These items can be viewed as benign tax avoidance. Tax aggressiveness can differ from tax avoidance in terms of scope as well; tax aggressive transactions (e.g., aggressive transfer-pricing policies) create tax avoidance, but tax avoidance effects are not necessarily due to tax aggressiveness (e.g., depreciation). Even within the realm of tax aggressiveness, there are degrees to which aggressiveness is relatively low (e.g., paying employee bonuses right before rather than right after fiscal year-end) or high (e.g., tax shelters). In this study, I focus on modeling the most extreme part of the tax minimization continuum—tax sheltering.

The recent archival literature has developed and used several measures for corporate tax avoidance and aggressiveness. First, Frank et al. (2009) measure tax aggressiveness as the residual from a regression of permanent book-tax differences (BTDs) on its known components, such as state tax expense, intangibles, and other items. Their rationale is that tax shelters, which are discretionary transactions, derive value as permanent BTDs that, by reducing current taxable income or generating tax credits, can understate the effective tax rate, boosting financial earnings.

Second, Dyreng et al. (2008) develop a measure of long-run tax avoidance as the ten-year sum of cash taxes paid divided by the ten-year sum of pretax income less special items. They appropriately note that tax avoidance might not pinpoint tax aggressiveness; avoidance can include legitimate deductions, while aggressiveness is a subset of avoidance because aggressiveness can also include prohibited activities such as tax sheltering. Nevertheless, several academic studies have used this measure as a proxy for tax aggressiveness.

Third, Mills (1998) finds that firms with greater BTDs are positively related to audit adjustments, implying that BTDs, at least in part, represent general tax risk. Prior literature also shows that BTDs capture financial earnings management (Hanlon 2005). To extend the analysis and examine whether information in BTDs are also informative of tax sheltering, I test for the relation between BTDs and tax shelters.

Finally, Gleason and Mills (2002) find that the contingent tax liability reserve, or tax cushion, is increasing in expected tax losses. One explanation is that a nontrivial portion of the tax cushion likely reflects tax sheltering transactions (also see Lisowsky et al. 2010).

To understand the extent to which various proxies for tax avoidance and tax aggressiveness measure tax sheltering activities in particular, the first supplemental analysis tests each of the above proxies' relation to actual tax shelter use.[9]

[9] Note that this analysis is not intended to find the most appropriate measure for general tax aggressiveness. Because tax aggressiveness is a matter of judgment, degree, and scope, some measures may not be related to tax sheltering in

Tax Sheltering in a FIN 48 Reporting Environment

The second supplemental analysis examines the usefulness of the model in inferring tax sheltering in a FIN 48 disclosure environment. For publicly traded firms with fiscal years beginning after December 15, 2006, FIN 48 requires disclosure of unrecognized (i.e., uncertain) tax benefits in the financial statement footnotes. A common conjecture is that FIN 48 provides the IRS with a "roadmap" with which to discover potentially aggressive tax positions (Mills et al. 2010). As such, the additional disclosures of tax uncertainty could curtail certain tax aggressive behavior, potentially weakening the applicability of the model in the FIN 48 environment. Therefore, I examine the model's usefulness in this new regime.

III. TAX SHELTERS AND THEIR CHARACTERISTICS

Tax Shelters

This study employs a new, expanded data set on corporate tax shelters obtained from the IRS Office of Tax Shelter Analysis (OTSA). These data capture 19 categories of specifically identified illegal corporate tax shelters known as "Listed Transactions," with interesting names such as Son of Boss (Notice 2000-44), Lease Strips (Rev. Rul. 2003-55), and Partnership Straddles (Notices 2002-50, 2002-65, and 2003-54).[10] These corporate tax shelters have been so widely used that the Treasury Department and IRS have issued Revenue Rulings, Treasury Regulations, and Notices explicitly prohibiting them.

Tax shelters are also grouped into a second category of (Non-Listed) "Reportable Transactions." These transactions are somewhat more general in nature.[11] The OTSA data used in this study capture 45 categories of corporate Reportable Transactions, such as corporate-owned life insurance (COLI), bank-owned life insurance (BOLI), and foreign sales corporation advance payments.[12] Listed or Reportable Transactions are reported for the fiscal year, similar to tax return filings (and financial statements). I define the indicator variable *TaxShelter* as equal to 1 if it is classified as a Listed or (Non-Listed) Reportable Transaction, and 0 otherwise. The data set covers the five-year period 2000 to 2004.

In general, Listed and Reportable Transactions achieve tax sheltering in one of three ways. First, they artificially inflate or shift asset bases to other (typically tax-indifferent) parties upon sale to decrease gains subject to tax or shift gains to a lightly taxed jurisdiction or party. Second, they accelerate or inflate expenses to decrease current taxable income. Third, they defer revenue to understate taxable income. Each of these ways can also influence the eligibility of certain tax credits that can be claimed directly against the firm's tax liability.

OTSA gathers data on tax shelters through both enforcement actions and voluntary compliance (i.e., self-reporting) by corporations. For example, IRS Form 8886 captures information on Listed and Reportable Transactions, such as their type, nature, and in some cases, dollar amount of

---

particular because they are not designed to do so. This result does not necessarily rule out their ability to capture other types of (non-tax shelter) tax aggressiveness or avoidance. The aim is to evaluate the ability of these various proxies to capture tax sheltering behavior in particular, that is, the most extreme case of tax minimization.

[10] The remaining 16 categories of corporate listed transactions used in this study are included in http:/www.irs.gov/businesses/corporations/article/0,,id=120633,00.html.

[11] To be more precise, Listed Transactions are a special class of Reportable Transactions. In particular, Listed Transactions are specifically identified tax shelters while the remaining (Non-Listed) Reportable Transactions are more general tax shelter transactions. For expositional purposes of this study, as well as to distinguish between the potentially different levels of tax sheltering, I separate the specific from general tax shelter transactions, calling the former "Listed" and the latter simply "Reportable." In effect, the short-hand term "Reportable" used in this study refers specifically to Reportable Transactions that are not already classified as Listed Transactions.

[12] Previous studies also interpret Reportable Transactions, such as COLI, as tax shelters (Graham and Tucker 2006; Brown 2009; Wilson 2009). Examples of other Reportable Transactions included in the OTSA data set are donations of intellectual assets, asset securitizations, and extraterritorial income exclusions.

tax benefits. After 2004, Schedule M-3 also captures some information on total Reportable (including Listed) Transactions, such as their book and tax income effects and a breakdown between temporary and permanent book-tax differences. This information is not available for my sample period, although I use both the 8886 and M-3 data later in the study to measure tax shelter activity in a more recent FIN 48 period.

Unfortunately, the OTSA data do not provide a clear indication as to the motivation for reporting. For example, corporate tax shelters thought to be "voluntarily" reported to the IRS could actually be due to enforcement actions or their threat. A related limitation is that *TaxShelter* can simply be a proxy for the probability of IRS audit. Although I cannot rule out this possibility, interdependencies surely exist between tax shelter reporting and audit regime, as one setting influences the other. I do not have data on IRS audit scoring models and therefore cannot control for this effect. However, the major benefit of using OTSA data is that inclusion in the sample is not conditional on whether the shelters are publicly disclosed, litigated, and/or settled. It provides the most concentrated sample of observations yet employed, increasing power of tests and as a result, generalizability for future research.

Tax Shelter Characteristics

Wilson (2009) finds that tax shelters are positively related to total book-tax differences, performance matched discretionary accruals (see Kothari et al. 2005), size, profitability (return on assets), and foreign income. I extend the Wilson (2009) model by also including constructs based on the conceptual framework from the Treasury (1999) detailing the characteristics of tax shelters. I focus on the additional Treasury (1999) characteristics here and refer the reader to Wilson (2009) for a discussion of the other characteristics.[13]

First, the Treasury (1999) explains that the use of tax-indifferent parties, such as foreign or tax exempt entities, is likely to be a characteristic of tax sheltering. Hines and Rice (1994) find that corporate subsidiaries located in tax haven or low-tax foreign jurisdictions can lower tax payments in general. Desai et al. (2006) explain that these entities can participate in transactions to absorb taxable income or deflect tax liability. As such, they find that tax haven operations facilitate a reduction in the tax burden by allowing firms to reallocate and defer the U.S. tax related to foreign income. Another method to shelter tax payments is by transacting with municipalities, as, for example, in the lease-in lease-out (LILO) transactions between Wachovia and U.S. and foreign municipalities (McGill and Outslay 2004).[14] I go a step further and test whether there is a positive relation between tax shelters in particular and tax haven operations, which can include low-tax or tax-indifferent subsidiaries.

To measure the tax haven aspect of tax-indifferent or low-tax counterparties, I develop an indicator variable called *TaxHaven*. Data are hand-collected from SEC Form 10-K Exhibit 21, which enumerates the state and country jurisdictions in which a U.S. company's subsidiaries are located. I cross-reference Exhibit 21 to the OECD (2000, 17) list of tax havens.[15] If a firm reports a subsidiary in such a tax haven, *TaxHaven* is set equal to 1, and 0 otherwise. Because locating a subsidiary in a tax haven is likely to be characteristic of tax sheltering consistent with the Treasury (1999), Desai et al. (2006), and Hines and Rice (1994), I predict a positive association between

[13] As in Wilson (2009), this study focuses on characteristics, not economic incentives, of tax shelter usage.

[14] Wachovia is not included in my sample. If it were, IRS data disclosure agreements would prevent me from identifying it. In the context of tax-indifferent parties, Wachovia simply illustrates how LILOs can be used in practice. LILO transactions in general are included in the OTSA sample data.

[15] The Organisation for Economic Co-operation and Development (OECD) tracks tax haven jurisdictions and urges their compliance with international standards on tax administration.

TaxHaven and tax shelters.[16]

Second, the Treasury (1999) explains that tax shelters exploit inconsistent financial and tax treatment of certain transactions. Therefore, transactions such as tax shelters that create permanent book-tax differences by lowering taxable income or generating tax credits, reduce a corporation's effective tax rate, increasing reported earnings and (perceived) shareholder value. The Treasury (1999, 7) notes that the focus on increasing corporate earnings and managing the reported effective tax rate has "caused many corporations to treat their in-house tax departments as profit centers," with corporate accountants viewing that "a high effective tax rate is a sign of weakness." McGill and Outslay (2002, 1136) explain that "most tax shelter products today have vitality because of their positive impact on the corporation's tax liability (reduced) and its earnings per share (increased)." Therefore, there is reason to expect that tax shelters are used to mitigate an otherwise high effective tax rate reported to shareholders.

I infer inconsistent book-tax treatment by examining two items. The first is the tax rate reported in the financial statements. The tax rate that is most directly related to earnings is the reported effective tax rate (ETR), or the ratio of total tax expense to total pretax income. The Treasury (1999) explains that the ETR is closely measured by corporate tax departments and can capture a key reporting incentive to engage in tax sheltering behavior, namely, to manage tax expense. Therefore, I predict that the greater the prior-year reported ETR (*LagETR*), or the ratio of total tax expense (Compustat #16) to total pretax income (#170), the more likely it is that a firm uses a tax shelter in the current year.[17]

The second item used in inferring inconsistent book-tax treatment is the indicator variable *EqEarn*, which is equal to 1 if equity method earnings (#55) are present, 0 otherwise. I measure equity method earnings because it is a peculiar case in consolidation accounting where a net percentage share of income and assets of a 20–49 percent-owned corporate investment is reported on a parent's consolidated financial statements, but remains on a cost basis on the consolidated tax return, where non-dividend income is not reported. Therefore, the equity method allows the firm to report greater net assets and income for financial purposes than tax purposes. I test whether such an arrangement is indicative of other advantageous consolidation arrangements, where the financial statements report greater profits than tax records. I predict a positive relation between tax shelters and *EqEarn*.

Third, the Treasury (1999) identifies financial engineering products as indicative of tax shelters. Certain hedges or structured transactions, for example, insulate a company from any significant economic risk or adverse financial statement reporting while engineering deductions or losses on the tax return (see Engel et al. 1999; Mills and Newberry 2005; Landsman et al. 2008). As the Treasury (1999) explains further, taxpayers can monetize the value in their assets or investments, or lock-in appreciation with respect to their property, without recognizing taxable gains (e.g., through borrowing).

To measure the use of financial engineering products, I use a proxy for mezzanine financing, *MezzFin*, capturing convertible debt and preferred stock (Compustat #39/#6). I use this variable because convertible debt has equity-like features, while preferred stock has debt-like features. Furthermore, Mills and Newberry (2005) explain (and empirically validate) that differences in

[16] Tax havens can also enable a firm to engage in other aggressive tax transactions, such as transfer-pricing or debt placement (interest stripping) schemes (see Wilson 1995; Mills et al. 1998).

[17] I use the prior-year rather than contemporaneous ETR because it is unclear the degree to which contemporaneous ETR is simultaneously related to tax shelters. Specifically, the contemporaneous ETR can either be the cause or effect of tax shelter use. I also use the reported ETR rather than the marginal tax rate (see Plesko 2003) because the focus of this item in the Treasury (1999) is on the financial reporting rather than pure economic issues of taxes. Nevertheless, sensitivity checks using contemporaneous ETR yield qualitatively similar results.

financial interest expense and tax interest deductions are associated with the advantageous classification of debt-equity hybrid securities for financial and tax reporting. I predict a positive relation between tax shelters and *MezzFin*.

Fourth, the Treasury (1999) (and U.S. Senate 2003) reports that the use of promoters, including investment banks, law firms, and auditors, is indicative of tax shelter usage. The white paper states that promoters attempt to design tax shelters so they can be replicated, marketed, and sold to a variety of clients. The Treasury (1999) reports that one Big 5 accounting firm maintained two databases of about 1,000 "mass market" tax savings ideas in the 1990s. Promoter usage is measured using *Big5*, equal to 1 if the firm is audited by PricewaterhouseCoopers, Ernst & Young, Deloitte & Touche, KPMG, or Arthur Andersen, and 0 otherwise (#149).[18]

Fifth, I examine more generally whether the firm's legal circumstances are informative in inferring tax shelters. A firm that is under lawsuit could be a harbinger of aggressive practices that also include tax sheltering, as is evident in the Graham and Tucker (2006) sample. I measure legal circumstances, *Litigation*, as equal to 1 if pretax (#372) or after-tax (#373) Litigation/Insurance Settlement is negative, i.e., a litigation or insurance payout. For both *Big5* and *Litigation*, I expect a positive relation to tax shelter usage.

Finally, I control for net operating loss carryforwards (*NOL*) of the firm. If the firm is already in a non-taxable position due to loss carryforwards, I predict a negative relation between tax shelters and *NOL*, equal to 1 if an NOL carryforward exists (#52 $>$ 0), 0 otherwise. Prior research suggests that loss firms have different characteristics than positive income firms (e.g., Joos and Plesko 2005; Klein and Marquardt 2006), and thus can affect the reliability of estimates reported in this study.[19]

IV. RESEARCH METHOD

Sample

The research sample begins with Compustat and Internal Revenue Service (IRS) corporate tax return data successfully matched on Employer Identification Number (EIN) and fiscal year-end from years 2000 to 2004. Only publicly traded Subchapter C corporations are included in this sample to ensure the company reports a tax position to the IRS.[20] This initial merged data set contains 26,388 firm-year observations (see Table 1).

I first restrict the initial data set to taxpayers with no foreign entity owning 25 percent or more of the corporation (from Form 1120, Schedule K, Line 7), leaving 16,776 firm-year observations. This step focuses the sample on U.S.-only domiciled firms by eliminating any foreign-controlled U.S. corporations. This elimination mitigates differences in international tax and financial reporting rules that can confound the results. Next, I require non-negative Implied Stock Option Expense reported in Compustat (#399), leaving 16,649 firm-year observations.[21] I then match and merge these observations to OTSA tax shelter data. OTSA only provides information on whether the firm uses a tax shelter, but does not provide confirmatory information on whether the firm does *not* use a tax shelter. Therefore, if confirmed identification of a tax shelter does not exist for a particular

[18] Arthur Andersen existed only for the earlier part of my sample period. Additional tests classifying the four currently operating accounting firms (PwC, E&Y, Deloitte, and KPMG) as promoters yield materially same results.

[19] In robustness checks, I also eliminate firms with negative pretax book income, yielding similar results.

[20] Subchapter S corporations, RICs, and REITs are excluded because they normally do not pay corporate income taxes. Subchapter C corporations are subject to the corporate income tax as calculated from the IRS Form 1120.

[21] In calculating the tax cushion later in the study, stock options are used to adjust for the difference between current U.S. tax expense and actual tax on the tax return. In such a construct for tax cushion, a negative stock option expense (i.e., stock option "income") is not usable. It implies a tax refund from stock option exercise, even though an investor would not exercise while in a loss (or "underwater") position.

TABLE 1

Sample Selection

Criteria	Firm-Year Observations: Compustat and IRS Data	Firm-Year Observations: OTSA Data[a]	Number of Tax Shelters (OTSA)
Subchapter C Corporations matched on EIN and Fiscal Year in Compustat and IRS tax return databases, 2000–2004.	26,388		
Less: Observations >25% foreign-owned, from Form 1120 Schedule K Line 7	(9,612)		
	16,776		
Less: Observations with negative Implied Option Expense (#399)	(127)		
	16,649	446	1,387
Less: Observations in Financial Industry (SIC 6000–6999)	(4,045)	(110)	(415)
	12,604	336	972
Less: Observations with negative current U.S. tax expense	(1,752)	(61)	(212)
	10,852	275	760
Less: Observations missing lagged values	(1,629)	(8)	(16)
Remaining 2000–2004 Sample Period Observations:	9,223	267	744
	[3,347 firms]	[211 firms][b]	

[a] OTSA (IRS Office of Tax Shelter Analysis) firm-year observations are a subset of the Compustat and IRS firm-year observations.

[b] Graham and Tucker (2006) sample: 43 unique firms with 152 total firm-year observations, 1975–2000.

Wilson (2009) sample: 59 unique firms with 215 total firm-year observations, 1975–2002.

Sample comparison: Wilson (2009) identifies 16 tax shelter firm-year observations in 2000–2002; the OTSA sample identifies 102 tax shelter firm-year observations in 2000–2002.

observation, I assume a tax shelter in fact does not exist, and therefore assign a zero value for that observation.

I delete observations in the financial industry (SIC 6000–6999) due to the general unavailability of data for my tests, and the unique nature of regulatory requirements, leaving 12,604 firm-year observations.[22] I further delete observations with negative current U.S. tax expense (Compustat #63; or Total Tax Expense [#16] less Deferred Tax Expense [#50]), leaving 10,852 observations. I delete these observations because, in calculating the tax cushion later in the study, the other input variables (actual tax on the tax return and the tax benefit from stock options) are zero or positive only. Finally, because lagged and change variables are required in the tax shelter likelihood model, I delete observations with missing prior-year data, leaving a final sample of

[22] In particular, values used in Wilson (2009) and inferred from the Treasury (1999), such as foreign pretax income, research and development, equity in earnings, and mezzanine financing are not available for financial firms. Also, leverage is influenced by regulatory requirements and can be otherwise incomparable with firms in other industries.

9,223 firm-year observations, of which 267 are tax shelter observations. Approximately one-third of the tax shelter observations are classified as Listed Transactions, while two-thirds are (Non-Listed) Reportable Transactions.

I do not employ a one-to-one matched sample in the main tests for two key reasons. First, from an out-of-sample estimation and generalizability perspective, it is unclear *a priori* which firms are shelter firms and which firms are not. Matching *a priori* on the item I ultimately want to infer, i.e., tax shelter usage, would limit the model's generalizability and potentially distort the extent of tax shelter representation in a larger population. Second, an unmatched sample allows the researcher to exploit greater variation in variables, increasing the potential to learn more specific details about the profile of tax shelter participants. However, to mitigate any concerns over my sample construction, I subsequently report model validity checks that also utilize matched samples derived from Graham and Tucker (2006).

Tax Shelter Likelihood Model

To generate a model using publicly available financial information to estimate the likelihood of tax shelters, I run the following pooled, cross-sectional logistic regression model by firm f in year t:

$$\ln \frac{P_{TaxShelter}}{1 - P_{TaxShelter}} = \alpha + \beta X + \varepsilon,$$

where $P_{TaxShelter} = \frac{1}{1+e^{-(\alpha+\beta X+\varepsilon)}}$ = probability the firm engages in a tax shelter transaction; and

$$\begin{aligned}\beta X = {} & \beta_1 BTD_{f,t} + \beta_2 DAP_{f,t} + \beta_3 Leverage_{f,t} + \beta_4 Size_{f,t} + \beta_5 ROA_{f,t} + \beta_6 ForeignIncome_{f,t} \\ & + \beta_7 R\&D_{f,t} + \beta_8 TaxHaven_{f,t} + \beta_9 LagETR_{f,t} + \beta_{10} EqEarn_{f,t} + \beta_{11} MezzFin_{f,t} + \beta_{12} Big5_{f,t} \\ & + \beta_{13} Litigation_{f,t} + \beta_{14} NOL_{f,t} + \sum_{y=15}^{18} \beta_y Year_{f,t} + \sum_{i=19}^{26} \beta_i Ind_{f,t}.\end{aligned}$$

See Table 2 for full explanation of variables. The first seven variables, book-tax differences (*BTD*), performance-matched discretionary accruals (*DAP*), leverage (*Leverage*), asset size (*Size*), profitability (*ROA*), foreign-source income (*ForeignIncome*), and research and development expense (*R&D*) are directly from Wilson (2009). The remaining seven variables are derived from the Treasury (1999), as described above. Industry variables control for variability in industry reporting. Consistent with Petersen (2009), I use Huber-White robust standard errors to adjust for potential serial correlation among multiple observations per firm, with year indicators controlling for cross-sectional correlation. All continuous variables are winsorized at the 1 and 99 percent levels to mitigate the effect of outliers.

V. RESULTS

Descriptive Statistics and Correlation Analysis

As detailed in Table 3, the sample is composed of 9,223 firm-year observations, of which 267, or 2.89 percent are tax shelter observations (*TaxShelter*; Panel A). About 8.65 percent of the sample reports a subsidiary in a tax haven jurisdiction (*TaxHaven*) while less than 12.5 percent reports equity method earnings (*EqEarn*). Almost two-thirds of the sample has a Big 5 auditor (*Big5*), and 4 percent reports a legal or insurance payout (*Litigation*). Over one-third of the sample reports a net operating loss carryforward (*NOL*). The most highly represented industries are Manufacturing (SIC 3000–3999), at 26 percent, and Chemical (SIC 2000–2999), at over 18 percent (Panel B). The least represented industries are Diversified (SIC 9000–9999), at 1.58 percent, and Agriculture (SIC 0000–0999), at 0.27 percent. Observations are generally evenly dispersed

TABLE 2

Variable Definitions

TaxShelter	= 1 if observation reports a Tax Shelter (Listed or Reportable) Transaction in the IRS Office of Tax Shelter Analysis (OTSA) database; 0 otherwise.
BTD	= (Pretax Book Income, #170, less Estimated Taxable Income)/Total Assets (#6), where Estimated Taxable Income = (current federal tax expense, #63, plus current foreign tax expense, #64)/0.35, less change in tax loss carryforward (#52).
DAP	= Discretionary accruals from the performance-adjusted modified cross-sectional Jones (1991) model (see Kothari et al. 2005; Wilson 2009).
Leverage	= Long-term debt liability (Compustat #9)/Total Assets (#6).
Size	= log(Total Assets, #6).
ROA	= Pretax Book Income (#170)/Total Assets (#6).
ForeignIncome	= Foreign Pretax Income (#273)/Lagged Total Assets (#6); 0 if #273 is missing.
R&D	= R&D Expense (Compustat #46)/Lagged Total Assets (#6); 0 if #46 is missing.
TaxHaven	= 1 if observation reports in 10K Schedule 21 a subsidiary located in a tax haven jurisdiction (0 otherwise), as defined by the OECD (2000): Andorra, Anguilla, Antigua and Barbuda, Aruba, Bahamas, Bahrain, Barbados, Belize, Bermuda, British Virgin Islands, Cayman Islands, Cook Islands, Cyprus, Dominica, Gibraltar, Grenada, Guernsey, Isle of Man, Jersey, Liberia, Liechtenstein, Maldives, Malta, Marshall Islands, Mauritius, Monaco, Montserrat, Nauru, Netherlands Antilles, Niue, Panama, Samoa, San Marino, Seychelles, St Lucia, St Kitts and Nevis, St Vincent and the Grenadines, Tonga, Turks and Caicos, U.S. Virgin Islands, Vanuatu.
LagETR	= Prior-Period Reported Effective Tax Rate $\times 100$, where Reported ETR = Total tax expense (Compustat #16)/Total Pretax Book Income (#170).
EqEarn	= 1 if Equity in Earnings (Loss) from income statement (#55) is present and does not equal 0, and 0 otherwise.
MezzFin	= Convertible Debt and Preferred Stock (Compustat #39)/Total Assets (#6); 0 if #39 is missing.
Big5	= 1 if Auditor is Deloitte & Touche, PricewaterhouseCoopers, Ernst & Young, KPMG, or Arthur Andersen (Compustat #149), and 0 otherwise.
Litigation	= 1 if Pretax (#372) or After-tax (#373) Litigation/Insurance Settlement is negative, indicating a payout, and 0 otherwise.
NOL	= 1 if Tax Loss Carryforward (#52) >0, and 0 otherwise.
Year	= 1 if observation falls in any of years 2001–2004, and 0 otherwise (if 2000).
Ind	= 1 if observation falls in any of one-digit SIC (non-financial) industries SIC0, SIC1, SIC2, SIC4, SIC5, SIC7, SIC8, and SIC9, and 0 otherwise (if SIC3). The financial industry (SIC6) is excluded from the sample.

through time, with slightly more observations in the earlier than later years. However, the uneven counts of *TaxShelter* in time and industry support inclusion of year and industry controls.

Examining the continuous variables for the sample, the mean (median) *BTD* is 21.1 percent (0.1 percent) of total assets. The mean (median) *DAP* is −0.010 (0.015) while *Leverage* has a mean (median) of 19.4 percent (7.3 percent). *Size* is highly variable, with mean (median) values of

TABLE 3

Descriptive Statistics: Tax Shelter Likelihood Model Variables

Panel A: Frequency of Indicator Variables

Indicator Variable	n	% of Sample	*TaxShelter* and Variable = 1	% of *TaxShelter*
TaxShelter	267	2.89%		100.00%
TaxHaven	798	8.65%	78	29.21%
EqEarn	1,149	12.46%	100	37.45%
Big5	6,087	66.00%	262	98.13%
Litigation	375	4.07%	31	11.61%
NOL	3,355	36.37%	74	27.72%

Panel B: Full-Sample Observations by Industry and Year

	2000	2001	2002	2003	2004	Total	% of Sample[a]	# of *TaxShelter*
sic0 (Agric.)	6	5	5	5	4	25	0.27%	<5
sic1 (Const.)	133	131	106	123	129	622	6.74%	21
sic2 (Chem.)	379	332	323	329	339	1,702	18.45%	50
sic3 (Mfg.)	664	488	414	437	431	2,434	26.39%	32
sic4 (Transp.)	208	190	178	180	191	947	10.27%	68
sic5 (Retl.)	297	251	232	231	248	1,259	13.65%	57
sic7 (Bus. Svc.)	344	313	257	275	286	1,475	15.99%	19
sic8 (Hlth.)	132	128	116	123	114	613	6.65%	17
sic9 (Div./Other)	33	27	27	32	27	146	1.58%	<5
Total	2,196	1,865	1,658	1,735	1,769	9,223		
% of Sample	23.81%	20.22%	17.98%	18.81%	19.18%		100.00%	
# of *TaxShelter*	38	46	18	86	79			267

(continued on next page)

Panel C: Distribution of Continuous Variables

Variable	25th Perc.	Median	Mean	75th Perc.	Std. Dev.
BTD	−0.050	0.001	0.211	0.226	2.198
DAP	−0.024	0.015	−0.010	0.136	0.360
Leverage	0.000	0.073	0.194	0.285	0.300
Size (log(Assets))	2.360	4.158	4.054	5.818	2.545
Assets ($M)	10.592	63.950	697.134	336.332	2,215.369
ROA	−0.242	0.022	−0.428	0.095	1.741
ForeignIncome	0.000	0.000	0.002	0.000	0.011
R&D	0.000	0.000	0.109	0.078	0.289
LagETR (%)	0.000	8.665	16.185	37.378	27.719
MezzFin	0.000	0.000	0.076	0.000	0.318

Panel D: Test of Differences in Means

	Means			
Variable	When *TaxShelter* = 1	When *TaxShelter* = 0	Difference in Means	t-statistic
BTD	0.046	−0.218	0.264	1.94*
DAP	0.066	−0.019	0.085	2.18**
Leverage	0.192	0.261	−0.069	−3.71***
Size (log(Assets))	7.804	3.942	3.862	25.26***
ROA	0.078	−0.443	0.521	4.82***
ForeignIncome	0.007	0.002	0.005	7.28***
R&D	0.016	0.112	−0.096	−5.34***
LagETR (%)	34.125	15.650	18.475	11.21***
MezzFin	0.019	0.078	−0.059	−2.95***

(continued on next page)

Panel E: Test of Differences in Proportions

Variable	Proportions When *TaxShelter* = 1	Proportions When *TaxShelter* = 0	Difference in Proportions	z-statistic
TaxHaven	29.21%	8.04%	21.17%	12.72***
EqEarn	37.45%	11.71%	25.74%	13.08***
Big5	98.13%	65.04%	33.09%	11.34***
Litigation	11.61%	3.84%	7.77%	6.61***
NOL	27.72%	36.63%	−8.91%	−2.99***

*,**,***Represent significance at the $p < 0.10$, 0.05, and 0.01 levels, respectively.

[a] Percentages do not equal 100.0% due to rounding.

See Table 2 for full variable definitions.

$697 ($64) million. The mean (median) profitability (*ROA*) is −42.8 percent (2.2 percent) of total assets. *ForeignIncome*, *R&D*, and *MezzFin* all report a zero median value, with respective means at 0.2 percent, 10.9 percent, and 7.6 percent. Finally, the lagged reported effective tax rate (*LagETR*) is 16.2 percent (8.7 percent), while the 75th percentile firm has a reported lagged ETR of over 37 percent (Panel C).

Tests for differences in means for users and non-users of tax shelters indicate that *BTD*, *DAP*, *Size*, *ROA*, *ForeignIncome*, and *LagETR* are all greater for tax shelter than for the non-tax shelter firms (all $p < 0.10$), consistent with prior models and expectations (Table 3, Panel D). The means tests show that *Leverage*, *R&D*, and *MezzFin* are significantly smaller for the tax shelter than for the non-tax shelter firms (all $p < 0.01$). The *Leverage* result is consistent with Graham and Tucker (2006), who find a negative relation between debt and tax shelter usage. *A priori*, it is unclear whether intangible activities facilitate tax sheltering, or like leverage, act as a substitute; the negative coefficient on *R&D* implies substitution. Finally, inconsistent with expectations, results show a negative relation between *TaxShelter* and mezzanine financing (*MezzFin*). It is likely that *MezzFin* (like *R&D*) substitutes for plain-vanilla *Leverage*, which itself is a substitute for tax shelters.

Tests for differences in proportions in Table 3, Panel E indicate that subsidiary locations in tax havens (*TaxHaven*), inconsistent book-tax treatment (*EqEarn*), use of promoters (*Big5*), and litigation payouts (*Litigation*) are significantly more likely for the tax shelter than non-tax shelter firms (all $p < 0.01$). Also, tax shelter firms are significantly less likely than non-shelter firms to report a net operating loss carryforward (*NOL*; $p < 0.01$).

Of the Wilson (2009) variables, untabulated Pearson and Spearman correlations show that *TaxShelter* is positively related to *BTD*, *Size*, *ROA*, and *ForeignIncome*, and negatively related to *Leverage* and *R&D* (all $p < 0.10$). *TaxShelter* is positively related to *DAP* only in Pearson correlations ($p < 0.05$). Of the Treasury (1999) variables, untabulated Pearson and Spearman correlations show that *TaxShelter* is positively related to *TaxHaven*, *LagETR*, *EqEarn*, *Big5*, and *Litigation* (all $p < 0.01$), and negatively related to *NOL* ($p < 0.01$). *MezzFin* is negatively related to *TaxShelter* in Pearson correlations only ($p < 0.01$).

Main Logistic Regression Results

This section reports the main efforts to use financial statement information in multivariate models to infer corporate tax shelter likelihood. To begin, the first column of Table 4 reports the findings of Wilson's Table 5, column 3. I compare and extend these results for two reasons. First, his sample composition is more than one-to-one matched, as is my OTSA-Compustat data. Second, this column's results are used later in his out-of-sample estimation and market tests. His tests show a positive relation between tax shelters and book-tax differences (*BTD*), discretionary accruals (*DAP*), size (*Size*), profitability (*ROA*), and *ForeignIncome*. Year and industry controls are not included. The Pseudo-R^2 is 13 percent.

The second column of Table 4 replicates the Wilson (2009) model on the OTSA data, again with no year or industry controls. Results are somewhat sharper using the OTSA data, corroborating the findings of Wilson (2009). In addition to his results, *Leverage* is now negatively related to *TaxShelter*, supporting Graham and Tucker (2006). Note, however, that *Leverage* calculated per Wilson (2009) only includes long-term debt (#9), while Graham and Tucker (2006) also include the short-term portion of long-term debt (#34). The Pseudo-R^2 of the Wilson (2009) model using OTSA data is 9.8 percent.

The third column of Table 4 replicates the Wilson (2009) model on the OTSA data, but adds year and industry controls. Results remain qualitatively identical as in the previous analysis, with an increase in the Pseudo-R^2 to 15.4 percent.

TABLE 4

Results: Tax Shelter Logistic Model

Dependent Variable: ***TaxShelter***

Variable	Pred. Coeff. Sign	Coefficients				
		Wilson as Reported[a]	**Wilson using OTSA Data (1)**	**Wilson using OTSA Data (2)**	**Combined Model (1)**	**Combined Model (2)**
Constant		−4.86 ***	−3.471***	−4.884***	−7.059***	−8.502***
		(−4.38)	(−29.20)	(−16.14)	(−8.13)	(−11.08)
BTD	+	5.20 *	0.756**	0.142***	0.032	0.136
		(1.54)	(2.13)	(3.14)	(0.60)	(1.25)
		[1.30]	[1.08]	[1.15]	[1.03]	[1.15]
DAP	+	4.08 *	0.748***	0.457**	0.924***	0.424
		(1.63)	(5.19)	(2.28)	(4.36)	(1.12)
		[1.02]	[2.11]	[1.58]	[2.52]	[1.53]
Leverage	−	−1.41	−1.036***	−0.963***	−0.697***	−0.671**
		(−1.18)	(−5.88)	(−5.25)	(−2.61)	(−2.40)
		[−0.35]	[0.36]	[0.38]	[0.50]	[0.51]
Size	+	0.76 ***	0.700***	0.728***	1.397***	1.311***
		(3.05)	(7.44)	(6.35)	(5.69)	(6.12)
		[0.19]	[2.01]	[2.07]	[4.04]	[3.71]
ROA	+	3.51 **	3.196***	3.235***	2.473***	2.425***
		(1.75)	(7.76)	(6.94)	(4.08)	(3.47)
		[0.88]	[24.44]	[25.42]	[11.86]	[11.30]
ForeignIncome	+	1.72 ***	17.305***	23.575***	3.569	8.673*
		(4.35)	(4.10)	(5.50)	(0.74)	(1.67)
		[0.43]	[>100.00]	[>100.00]	[35.46]	[>100.00]
R&D	+	2.43	−4.391	−2.359	−3.023	−1.026
	or	(0.58)	(−1.17)	(−0.99)	(−0.88)	(−0.62)
	−	[0.61]	[0.01]	[0.09]	[0.049]	[0.358]
TaxHaven	+				0.790***	1.039***
					(4.46)	(5.54)
					[2.20]	[2.83]
LagETR	+				0.015***	0.017***
					(4.96)	(5.67)
					[1.02]	[1.02]
EqEarn	+				1.048***	0.969***
					(6.32)	(5.69)
					[2.85]	[2.64]
MezzFin	+				−0.687	−0.731
					(−1.07)	(−1.19)
					[0.50]	[0.48]
Big5	+				3.186***	3.008***
					(3.45)	(3.82)
					[24.20]	[20.24]
Litigation	+				1.063***	0.952***
					(4.84)	(4.15)
					[2.89]	[2.59]
NOL	−				0.140	0.100

(continued on next page)

TABLE 4 (continued)

Dependent Variable: ***TaxShelter***

Variable	Pred. Coeff. Sign	Coefficients: Wilson as Reported[a]	Wilson using OTSA Data (1)	Wilson using OTSA Data (2)	Combined Model (1)	Combined Model (2)
					(0.81)	(0.54)
					[1.15]	[1.11]
Industry Controls		Not Included	Not Included	Included	Not Included	Included
Year Controls		Not Included	Not Included	Included	Not Included	Included
Observations		319	9,223	9,223	9,223	9,223
Pseudo-R^2		13%	9.8%	15.4%	20.4%	25.8%

*,**,***Represent significance levels at the p <0.10, 0.05, and 0.01 levels, respectively.

[a] Results are as reported in Wilson (2009, Table 5, column 3).

z-statistic (odds ratio) appears in parentheses (brackets) below coefficient estimate, except the bracketed terms in the Wilson as Reported column indicate the marginal change in the probability of using a tax shelter resulting from a change in the independent variable. Huber-White standard errors are used to adjust for potential serial correlation among multiple observations per firm. Continuous variables are winsorized at the 1 and 99 percentile levels.

See Table 2 for full variable definitions.

All models reported herein are significant at the p <1 percent level using a Chi-square test.

The fourth column of Table 4, Combined Model (1), combines the Wilson (2009) variables with the Treasury (1999) variables developed in this study. In this column, no year or industry controls are included to provide comparability between models. Combined Model (1) reports that *BTD* and *ForeignIncome* are no longer positively related to *TaxShelter*, whereas *DAP*, *Leverage*, *Size*, and *ROA* remain significant with the expected signs. Several Treasury (1999) variables are significant as well. Results show that *TaxShelter* is positively related to the tax-indifferent party proxy *TaxHaven*, the lagged effective tax rate *LagETR*, inconsistent book-tax treatment *EqEarn*, use of promoters *Big5*, and *Litigation*. *MezzFin* and *NOL* are not significantly related to *TaxShelter*. The Pseudo-R^2 rises to 20.4 percent.

The final column of Table 4, Combined Model (2), replicates the preceding analysis, but includes year and industry controls. The Wilson (2009) variables *BTD* and *DAP* are not significant. *Size*, *ROA*, and *ForeignIncome* are positively related to *TaxShelter*, while *Leverage* is negatively related. The Treasury (1999) variables *TaxHaven*, *LagETR*, *EqEarn*, *Big5*, and *Litigation* remain positively related to *TaxShelter* (all $p < 0.01$). The Pseudo-R^2 of this model is 25.8 percent.

These results extend the original Wilson (2009) model by including proxies for tax shelter characteristics consistent with the Treasury (1999). The result on *TaxHaven* adds to the findings of other studies, such as Desai et al. (2006), suggesting that tax havens and tax-indifferent parties facilitate the most severe form of tax aggressiveness. Indeed, examining the OTSA sample more closely, no less than 107 observations have tax shelter transactions involving some form of tax-indifferent parties. Reportable Transactions included in this number are as a result of foreign sales corporations, foreign tax credits, extraterritorial income exclusions, and foreign worthless stock (n = 30). The largest proportion of Listed Transactions is LILO (n = 68). The remaining observations are accounted for by the Listed Transactions Inflated-Basis CARDS and *ACM Partnership v. Commissioner* (n = 9).

Next, the positive relation between the lagged reported effective tax rate (*LagETR*) and the likelihood of engaging in a tax shelter supports the idea that tax sheltering is higher in firms with relatively larger past reported tax rates. This finding is consistent with anecdotal evidence in the Treasury (1999) of pressure on tax departments to act like a "profit center," engineering higher earnings. It is also consistent with firms resorting to tax shelters if other, less aggressive methods are unavailable. The positive sign on the inconsistent book-tax treatment variable *EqEarn* indicates that the likelihood of tax shelter usage increases in book-tax consolidation differences. Such treatment of entities or transactions could be indicative of advantageous reporting in the financial statements relative to tax records.

The financial complexity variable, *MezzFin*, is not incrementally significant in the main tests. Although debt-equity hybrid instruments can be related to more general tax avoidance, results here indicate that they are not likely related to tax sheltering behavior in particular, after controlling for other measured characteristics.

The positive relation between *TaxShelter* and *Big5* lends some support to the idea that tax shelter use is increasing in the presence of (and access to) promoters. The positive sign on the legal circumstances variable *Litigation* indicates that insurance or litigation payouts are an informative sign of tax sheltering. Specifically, firms that are found to be aggressive from a legal standpoint seem to be aggressive from a tax standpoint as well.

Note that the odds ratios reported below each coefficient and z-statistic represent the odds of the given variable successfully identifying tax shelters to the odds it does not. An odds ratio of 1 indicates no discriminatory power (equal chance of the variable identifying tax shelters). In the last column of Table 4, for example *TaxHaven* (*Litigation*) reports an odds ratio of 2.83 (2.59), indicating that a firm with a subsidiary in a tax haven (litigation payouts) is 2.83 (2.59) times more likely to use a tax shelter than not.

In all, the Treasury (1999) white paper provides a useful conceptual guide in identifying an extended set of predictors that augment Wilson's (2009) analysis. The logistic model finds support that tax shelters are positively associated with tax haven subsidiaries, lagged reported effective tax rates, inconsistent book-tax treatment, the use of promoters, and litigation payouts, as well as size, profitability, and foreign income. Tax shelter usage is also negatively related to leverage, indicating a substitution effect. After controlling for the above relations, it does not appear that there is a strong incremental positive link between tax shelters and book-tax differences, discretionary accruals, or mezzanine financing.

Out-of-Sample Validation Tests

To examine the validity of Table 4's Combined Model (2), I perform several validation tests. The first two tests run this model on matched, out-of-sample period tax shelter observations. The third test calculates the tax shelter likelihood model on the first four years of my sample (2000–2003), then applies the parameter estimates to a "holdout" sample from year 2004. I also report Receiver Operator Characteristic (ROC) Curve results from each of these tests. In brief, the validation tests indicate that the main model is significantly helpful in both inferring and predicting tax shelter likelihood.

First, I identify out-of-sample tax shelter observations using Graham and Tucker (2006). I limit observations to the 1994–1998 period due to limitations in gathering *TaxHaven*.[23] Also, although Graham and Tucker (2006) include financial institutions in their study, I exclude them due to data availability issues on several variables as described above. This out-of-sample group

[23] SEC Forms 10-K (and Schedules 21) are available online from 1994 onward, making hand-collection of the *TaxHaven* variable more feasible in this later period.

yields 21 usable tax shelter firm-year observations, to which I match a control sample of 21 non-tax shelter observations.[24] If an observation is a tax shelter observation, the indicator variable *G&TTaxShelter* is set equal to 1, or 0 otherwise. I use a one-to-one match to validate whether the original model will apply to such a sample.

For this pooled sample of 42 observations, I calculate a tax shelter score (*TaxShelterScore1*) using firm-specific financial information as inputs into the reported coefficients of Combined Model (2) in Table 4. I then run a logistic regression of *G&TTaxShelter* on *TaxShelterScore1* to examine if the score is positively related to out-of-sample tax shelter likelihood. Results in Table 5, Panel A show that *TaxShelterScore1* is positively related to *G&TTaxShelter* ($p < 0.05$), indicating the model's parameters are useful in identifying out-of-sample tax shelters, even in a one-to-one matched sample setting.

To supplement this test, I also calculate the area under the ROC Curve (see Hosmer and Lemeshow 2000). An area under the ROC Curve of 0.50 represents a model that does not have discriminatory power beyond that of chance, while higher levels represent improved power. The ROC Curve has been used extensively in the bankruptcy literature (e.g., Chava and Jarrow 2004) and accounting literature (e.g., Demers and Joos 2007) to estimate both the validity and predictive ability of logistic models on out-of-sample period observations. Results on the matched sample show that the area under the ROC Curve is approximately 0.71, indicating strong discriminatory power of the model to identify tax shelter firms. In short, for a random pair of firms with and without tax shelters, the model's inferred predicted probability of tax shelter usage (i.e., *TaxShelterScore*) correctly identifies the tax shelter firm approximately 71 percent of the time.

I repeat the analysis using the Graham and Tucker (2006) observations, but match ten control observations, randomly picked from Compustat, to every one tax shelter observation. The only constraints I apply are a same-year match and that total assets (#6) are available. I again calculate a tax shelter score (*TaxShelterScore2*) by applying the parameters calculated from Table 4 to each observation. Results in Table 5, Panel B indicate a significantly positive relation between *TaxShelterScore2* and *G&TTaxShelter* ($p < 0.01$). ROC analysis reveals that the area under the curve is 0.69, indicating strong discriminatory power.

As the Wilson (2009) replication tests suggest, there likely are differences in how the Graham and Tucker (2006) and OTSA tax shelter observations are gathered, and the time period 1994–1998 might be significantly different from 2000–2004. Because of these potential differences, I re-estimate the logistic model on observations in years 2000 through 2003, leaving 2004 as the "holdout" sample period. For each observation in 2004, I calculate its tax shelter score (*TaxShelterScore3*) using the estimated parameters from the model estimated on the 2000–2003 observations (untabulated). The logistic regression reveals a strong positive relation between *TaxShelter*, as originally defined, and *TaxShelterScore3* for year 2004 observations ($p < 0.01$; Table 5, Panel C). ROC results show that the area under the curve for this holdout sample is 0.88, indicating an 88 percent chance that a firm with tax shelters has a higher predicted probability score (i.e., *TaxShelterScore3*) than a firm without tax shelters, for random pairs of firms with and without shelters.[25]

In addition to the ROC Curves, Table 5, Panel D reports the ability of the various "tax shelter scores" to identify tax shelters within 10-percentile increments in the distribution of

[24] Observations are matched on same-year, same one-digit SIC code, availability of total assets (#6), and closest market-to-book (MTB) ratio (number of common shares outstanding [#25] × price per common share at fiscal year-end [#199], divided by book value of common equity [#60]). In untabulated tests, MTB is not related to *TaxShelter*.

[25] As a comparison, bankruptcy models' ROC Curve areas range from 0.74 to 0.91 (Chava and Jarrow 2004).

TABLE 5

Validation Results

Panel A: Out-of-Sample Tax Shelter Model Results (1:1 Match on Industry, Year, and MTB)[a]

Dependent Variable: *G&TTaxShelter*

Variable	Pred. Coeff. Sign	Coefficient Using Full Model
Constant		−0.911*
		(−1.84)
TaxShelterScore1	+	2.016**
		(2.48)
Model Chi-square		6.87***
Observations		42
Pseudo-R^2		11.8%
Area under ROC Curve		0.71

Panel B: Out-of-Sample Tax Shelter Model Results (10:1 Random Match on Year)[b]

Dependent Variable: *G&TTaxShelter*

Variable	Pred. Coeff. Sign	Coefficient Using Full Model
Constant		−3.099***
		(−7.67)
TaxShelterScore2	+	1.675***
		(3.08)
Model Chi-square		10.18***
Observations		231
Pseudo-R^2		7.2%
Area under ROC Curve		0.69

Panel C: Hold-Out Sample Tax Shelter Model Results (Year 2004 OTSA Data)[c]

Dependent Variable: *TaxShelter*

Variable	Pred. Coeff. Sign	Coefficient Using Full Model
Constant		−6.238***
		(−8.32)
TaxShelterScore3	+	3.832***
		(4.98)
Model Chi-square		79.32***
Observations		1,769
Pseudo-R^2		12.3%
Area Under ROC Curve		0.88

(continued on next page)

Panel D: Distribution Analysis of *TaxShelterScores*[d]

	Cumulative Percent of Tax Shelters Correctly Identified		
Percentile of Sample	**1:1 Match 21 Tax Shelter Obs.**	**10:1 Match 21 Tax Shelter Obs.**	**2004 Holdout Sample 79 Tax Shelter Obs.**
91–100	10%	5%	52%
81–90	24%	29%	80%
71–80	38%	57%	89%
61–70	57%	76%	95%
51–60	67%	86%	95%
41–50	76%	86%	97%
31–40	86%	90%	99%
21–30	90%	90%	99%
11–20	95%	100%	99%
0–10	100%	100%	100%
Total Observations	42	231	1,769

*, **, *** Represent significance levels at the $p < 0.10$, 0.05, and 0.01 levels, respectively. z-statistic appears in parentheses below coefficient estimate.

[a] *G&TTaxShelter* equals 1 if Graham and Tucker (2006) identify observation as reporting a tax shelter, and 0 otherwise. *TaxShelterScore1* is the score for this sample using the Combined Model (2) estimation equation and coefficients reported in Table 4. Half of observations are Tax Shelter observations identified from Graham and Tucker (2006, 1994–1998, i.e., *G&TTaxShelter* = 1. The other half is a sample of control observations matched to tax shelter observations based on year, industry, market-to book ratio (Compustat #25 number of common shares outstanding ×Compustat #199 price per common share at end of fiscal year, divided by Compustat #60, book value of common equity), and availability of total assets (#6), i.e., *G&TTaxShelter* = 0.

[b] *G&TTaxShelter* equals 1 if Graham and Tucker (2006) identify observation as reporting a tax shelter, and 0 otherwise. *TaxShelterScore2* is the score for this sample using the Combined Model (2) estimation equation and coefficients reported in Table 4. Tax shelter observations are identified from Graham and Tucker (2006), 1994–1998. To each of the tax shelter observations identified in Graham and Tucker (2006), i.e., *G&TTaxShelter* = 1, 10 observations are matched randomly according to tax shelter year and availability of total assets (#6), i.e., *G&TTaxShelter* = 0.

[c] *TaxShelter* equals 1 if observation reports a Tax Shelter (Listed or Reportable) Transaction in the IRS Office of Tax Shelter Analysis (OTSA) database, and 0 otherwise. The tax shelter likelihood model is first estimated on observations in years 2000 through 2003. Then the coefficients are applied to year 2004 observations (holdout sample) to calculate *TaxShelterScore3*.

[d] Tax shelter scores are inferred from Table 4 and Table 5 parameters (i.e., *TaxShelterScore1* for 1:1 Matched sample; *TaxShelterScore2* for 10:1 Matched Sample; and *TaxShelterScore3* for the 2004 Holdout Sample). The ranking is ordered from highest tax shelter scores in the 91–100th percentile to the lowest in the 0–10th.

observations.[26] Here, tax shelter scores are sorted and ranked from highest (91–100th percentile) to lowest (0–10th percentile) for each sample.[27] In the one-to-one matched Graham and Tucker (2006) sample, the resulting firm-specific tax shelter score (*TaxShelterScore1*) clusters two-thirds of the tax shelter users in the top half of sorted observations. In the ten-to-one matched sample, the resulting tax shelter score (*TaxShelterScore2*) identifies the majority of tax shelters in the top 30 percent of sorted observations. When the (2000–2003) model is applied to the 2004 holdout sample, the tax shelter score (*TaxShelterScore3*) identifies the majority of tax shelters in the highest ten percent of sorted observations.

[26] This panel is essentially a tabular summary of the ROC Curve, but it is included to provide slightly greater detail.

[27] This method mimics the IRS audit selection process, in which tax compliance risk is quantified and ranked from highest to lowest for taxpayers. Such a method increases the probability of selecting highest-risk taxpayers to audit.

In sum, the tests of the likelihood model on out-of-sample and holdout observations lend support that the models developed it this study are largely valid and can therefore be used by investors, analysts, tax administrators, and tax and non-tax researchers to generate point estimates of corporate tax shelter likelihood.

VI. SUPPLEMENTAL ANALYSES

Linking Tax Shelters to Alternative Proxies of Tax Aggressiveness

The main focus of this study has been to extend and validate the tax shelter likelihood model using variables in Wilson (2009) and the Treasury (1999). Supplemental analyses examine the link between tax shelter use and various proxies of tax avoidance and tax aggressiveness, and compare these to the model developed in this study. These tests aim to shed light on the idea of a tax minimization continuum. To examine the relation of each proxy to tax shelters, I replicate the approach of the holdout sample analysis in Table 5, Panel C. In particular, I estimate a tax shelter likelihood model (i.e., a univariate logistic regression of *TaxShelter* on the tax aggressiveness proxy of choice) on 2000–2003 firm-year observations, then apply the parameters to actual data in 2004.

I examine four widely used proxies for tax avoidance and tax aggressiveness. The first is from Frank et al. (2009), who calculate tax aggressiveness as the residual from a regression of permanent book-tax differences on their known determinants of intangibles, unconsolidated subsidiaries, minority interest, current state tax expense, change in NOL, and lagged permanent book-tax differences. Summary results from the holdout logistic regression show that the discretionary permanent book-tax differences are not significantly discriminatory in inferring actual tax shelter usage ($\chi^2 = 0.15$, $p > 0.10$), with the area under the ROC Curve being 0.58 (see Table 6).

One potential explanation for the lack of results is that the Frank et al. (2009) measure is based on a broader effective tax rate differential, which includes more than permanent book-tax income differences, such as tax credits (Hanlon and Heitzman 2009), while tax shelters have also been found to be related to temporary book-tax differences (Wilson 2009; Lisowsky et al. 2010). Therefore, perhaps the Frank et al. (2009) measure captures tax aggressiveness along other portions of the tax minimization continuum not tested here. However, because they also validate their measure on tax shelter observations from Graham and Tucker (2006), the results here should be interpreted with caution.[28]

Next, I examine the long-run cash effective tax rate (ETR; Dyreng et al. 2008). I calculate this value as the ten-year sum of cash taxes paid, divided by the ten-year sum of pretax income less special items.[29] As expected, this sample size is the smallest due to a requirement for adequate time-series data. Results show that the long-run cash ETR is not significantly discriminatory in inferring tax shelter use ($\chi^2 = 1.68$, $p > 0.10$). The area under the ROC Curve is 0.58. This result supports the observation in Dyreng et al. (2008) that their measure is intended to capture tax avoidance generally and might not necessarily capture specific forms of tax sheltering. The results

[28] In further examining the link between financial aggressiveness and tax shelters, I test two separate models with *Listed* (*Reportable*) as the dependent variable, equal to 1 if the tax shelter is a Listed (Reportable) Transaction, 0 otherwise. In models without controls, discretionary accruals (*DAP*) are positively related to *Listed* and *Reportable* (both $p < 0.10$). In models with controls, *DAP* is not related to *Listed* or *Reportable*. Although the link between financial aggressiveness and tax shelters appears tenuous here, these tests also assume that *DAP* is a powerful proxy for financial aggressiveness. See Wysocki (2004) for a discussion on the validity of current discretionary accruals models capturing financial aggressiveness, and on the potentially endogenous nature of earnings management and tax compliance. Robustness checks using other measures for financial aggressiveness do not find a relation between *TaxShelter* and firms that (1) re-state their income tax expense; (2) re-state their net income; or (3) are issued an Accounting and Auditing Enforcement Release (AAER) from the SEC.

[29] I also use the five-year long-run cash ETR, with qualitatively identical results.

TABLE 6

Summary Results of Extended ROC Curve Analysis Linking Tax Shelters to Various Proxies for Tax Aggressiveness

Discretionary Permanent Book-Tax Differences	**Long-Run Cash ETR**
n = 801	n = 626
Wald Chi-square: 0.15 (p >0.10)	Wald Chi-square: 1.68 (p >0.10)
Area under ROC Curve: 0.58	Area under ROC Curve: 0.58
Total Book-Tax Differences	**Tax Cushion (Using IRS Data)**
n = 1,769	n = 1,769
Chi-square: 34.18 (p <0.01)	Chi-square: 40.98 (p <0.01)
Area under ROC Curve: 0.61	Area under ROC Curve: 0.69

Similar to Table 5, Panel C, individual regression equations of *TaxShelter* are first run for years 2000–2003 on each of the above proxies for tax aggressiveness. After obtaining coefficient estimates, the respective actual tax aggressiveness proxy calculated as of 2004 is input into its related estimation equation. The area under the ROC Curve summarizes the ability of each tax aggressiveness variables to infer actual *TaxShelter* in 2004, given its relation to *TaxShelter* in 2000–2003. Discretionary Permanent Book-Tax Differences are calculated as the residual from a regression of Permanent *BTDs* (*BTD* as calculated in this study less Temporary *BTDs*, or Deferred Tax Expense (#50)/0.35) on Intangibles (#33/#6), Unconsolidated Subsidiaries (#55/#6), Minority Interest Income (Loss) (#49/#6), Current Income State Tax Expense (#173/#6), Change in *NOL* (#52/#6), and Lagged Permanent *BTDs* (from Frank et al. 2009). The Long-Run Cash Effective Tax Rate (*ETR*) is calculated as the ten-year sum of cash taxes paid (#317) divided by the ten-year sum of pretax income (#170) less special items (#17) (from Dyreng et al. 2008). Results are qualitatively identical using the five-year sum. Total Book-Tax Differences are calculated as in this study, *BTD*. Tax Cushion is calculated as follows: {Current U.S. Tax Expense (#63) [or Total Tax Expense (#16) minus Deferred Tax Expense (#50) if #63 is 0] minus Tax Benefit from Stock Options [(#399/0.65) * 0.35] minus Total Tax After Credits (IRS Form 1120 page 1, line 31)} divided by Total Assets (#6).

using the Frank et al. (2009) and Dyreng et al. (2008) measures support the idea that tax avoidance, aggressiveness, and sheltering proxies likely capture unique portions along the tax minimization continuum.

Third, I examine total book-tax differences, calculated as *BTD* in this study, or the difference between pretax book income and estimated taxable income, scaled by total assets. Results show that total book-tax differences alone are helpful in gleaning actual tax shelter behavior ($\chi^2 = 34.18$, $p < 0.01$). However, the area under the ROC Curve is 0.61, indicating that book-tax differences can correctly discriminate between a tax shelter and non-tax shelter firm only 61 percent of the time. One explanation for the weaker result is that a firm can in fact have a contemporaneously low (or negative) level of book-tax differences, then engage in a tax shelter without increasing its *BTD* to a high level, in absolute terms. Nevertheless, *BTD alone* appears to be a crude, yet informative (and easily implementable) measure for corporate tax sheltering. However, recall that the main multivariate tests show that *BTD* is not related to *TaxShelter* when controlling for other firm characteristics.

Finally, I test the link between tax shelter usage and the current-year accrual of the contingent tax liability reserve, or tax cushion. For this test I employ tax return data, calculating tax cushion as the difference between current U.S. tax expense and Total Tax after Credits (IRS Form 1120 Page 1 Line 31), adjusted for the tax benefit of stock options ([#399/0.65] × 0.35). I test the discriminatory power of tax cushion in order to evaluate the extent to which the IRS has superior information than the public in inferring tax risk. Results show that for the holdout sample, tax cushion is significantly related to actual tax shelter use ($\chi^2 = 40.98$, $p < 0.01$). The area under the ROC Curve is 0.69. This result suggests that a nontrivial portion of the tax cushion is linked to tax

shelters. This result also suggests that tax shelter firms correctly anticipated the possibility of tax shelters being discovered and eventually disallowed by tax administrators, even before FIN 48.

Overall, the univariate holdout tests between tax shelters and alternative proxies of tax avoidance and aggressiveness show that total book-tax differences (BTDs) and the current-year tax cushion accrual are useful in inferring tax shelters, but discretionary permanent BTDs and long-run effective cash ETRs are not. However, *TaxShelterScore* retains the most significant discriminatory power relative to the other measures, with the main model's ROC Curve area of 0.88. This result supports the refinements of the Treasury (1999) to the Wilson (2009) model and provides researchers, investors, and policymakers with an additional tool with which to infer corporate tax sheltering behavior.

Inferring Tax Sheltering in a FIN 48 Environment

The second supplemental test examines the main model's usefulness to infer tax shelters during the FIN 48 regime. The assumption is that, when firms are compelled to make greater disclosure of their uncertain tax benefits, they are less likely to engage in tax aggressive—including tax sheltering—behavior, thus reducing the power of the model. To evaluate this possibility, I use a separate sample of 687 publicly traded firms with December 31, 2007, fiscal year-end dates that are part of the S&P 1500, with available financial data to calculate the respective *TaxShelterScore* (see Lisowsky et al. [2010] for a similar sampling methodology). I calculate *TaxShelterScore4* using the reported coefficients in Table 4's Combined Model (2) by inputting the 2007 firm data into this model's parameters.

To measure tax shelter usage in 2007, I obtain additional IRS data on Listed and (Non-Listed) Reportable Transactions. I set *TaxShelter_FIN48Period* equal to 1 if in 2007 a firm reports Reportable Transactions on Schedule M-3 Line 12 ("Items related to reportable transactions"), or reports a Listed or (Non-Listed) Reportable Transaction on Form 8886, Line 2 ("Type of reportable transaction"). I then run a logistic regression of *TaxShelter_FIN48Period* on *TaxShelterScore4*. Results show a significant, but weaker ability to correctly classify tax shelter firms than in Table 4, Combined Model (2). First, *TaxShelterScore4* is marginally significant at the 10 percent level. Also, the area under the ROC Curve is only 0.61 (see Table 7). Although still significant, it appears that the ability of Combined Model (2) to infer tax sheltering is more muted in a FIN 48 than in a pre-FIN 48 reporting environment. These results suggest that tax shelter behavior dissipates because of greater financial disclosure requirements (see Mills et al. 2010). However, the weaker results can also be due to lower power, e.g., different and fewer observations in 2007 compared to 2000–04; and/or the 2007 sample does not capture potential front-loading of tax benefits if the initial tax shelter year is outside of the sample period (e.g., LILO). Also, other clues in the FIN 48 disclosures could be more informative to investors than the model (Frischmann et al. 2008; Lisowsky et al. 2010). Further research is needed to analyze these explanations. As a result, the main model should be used with caution in the recent FIN 48 environment.

VII. CONCLUSION

This study extends prior empirical work on inferring tax shelter participation (Wilson 2009) by applying a conceptual framework from the Treasury (1999) white paper on tax shelters. In doing so, I employ confidential tax shelter data obtained from the IRS, increasing sample sizes, power, and generalizability. Main results show that the likelihood that a firm uses tax shelters is positively related to the presence of its subsidiaries located in tax havens, foreign-source income, inconsistent book-tax treatment, litigation losses, use of promoters, profitability, and size; and is negatively related to leverage. The model developed in this study increases the explanatory power of the Wilson (2009) model and corroborates the negative relation between debt and tax shelters found in Graham and Tucker (2006). I also develop a metric called *TaxShelterScore*, which

TABLE 7

Inferring Tax Sheltering in a FIN 48 Environment
Out-of-Sample Tax Shelter Model Results

Dependent Variable: ***TaxShelter_FIN48Period***

Variable	Pred. Coeff. Sign	Coefficient Using Full Model
Constant		−1.230*
		(−5.32)
TaxShelterScore4	+	0.380*
		(1.78)
Model Chi-square		24.95***
Observations		687
Pseudo-R^2		3.4%
Area under ROC Curve		0.61

*, **, *** Represent significance levels at the $p < 0.10$, 0.05, and 0.01 levels, respectively. z-statistic appears in parentheses below coefficient estimate.

TaxShelter_FIN48Period equals 1 if a firm reports a listed or reportable transaction on its IRS Schedule M-3 or Form 8886 in 2007, and 0 otherwise. *TaxShelterScore4* is the score for this sample using the Combined Model (2) estimation equation and coefficients reported in Table 4. The observations (n = 687) were chosen from calendar year-end publicly traded taxpayers part of the S&P1500 for which information to calculate necessary variables were available. See Lisowsky et al. (2010) for a similar sample selection composition.

estimates the probability that a firm engages in a tax shelter. Tests using *TaxShelterScore* show that the models have strong discriminatory power in inferring corporate tax shelter use. In applying *TaxShelterScore* to a holdout sample of firms, the area under the ROC Curve is 0.88, indicating an 88 percent chance that a firm with tax shelters has a higher predicted probability (or *TaxShelterScore*) than a firm without a tax shelter.

In additional analyses, I examine the relative benefits of the main model's refinements in measuring tax shelters by testing the link between tax shelters and four other proxies for tax avoidance and tax aggressiveness. Results show a strong positive relation between tax shelter usage, and both total book-tax differences (Mills 1998) and the contingent tax liability reserve, or tax cushion (Gleason and Mills 2002). However, there is no significant association between tax shelter usage and either discretionary permanent book-tax differences (Frank et al. 2009) or long-run cash effective tax rates (Dyreng et al. 2008). The main model remains the most highly discriminatory. The results shed light on the extent to which various measures do or do not capture tax sheltering. Such insight is helpful to researchers in designing appropriate measurements and tests when the degree of tax avoidance, aggressiveness, or sheltering, is of paramount interest.

Finally, I test the usefulness of the model in the new disclosure regime of FIN 48, in which firms must disclose unrecognized tax benefits. Results show that the *TaxShelterScore* metric is weaker, but remains significant, in identifying tax sheltering during this time. Therefore, inferences made using this model in a FIN 48 period should be made with care.

As a final note of caution, because the model developed in this study requires a variety of input variables, and is designed to measure the most extreme type of tax aggressiveness—tax shelters—it can either limit or sharpen the types of research questions for which the model is well-suited. Therefore, the model can potentially aid researchers in further exploring interesting questions on tax sheltering in particular, such as its relation to managerial compensation, value creation, and financing risks via the cost of capital. Further, U.S. and international financial and

tax policymakers would be interested in these results to the extent changes in reporting and disclosure practices are desirable. For example, future research could examine whether book-tax conformity reduces (or exacerbates) tax sheltering, and whether tax sheltering behavior changes with additional oversight (e.g., corporate governance) or audit demands (e.g., Sarbanes-Oxley).

REFERENCES

Boynton, C., P. Lisowsky, and W. Trautman. 2008. E-file, enterprise structures, and tax compliance risk assessment. *Tax Notes* (September 15): 1069–1078.

Brown, J. 2009. The spread of aggressive corporate tax reporting: a detailed examination of the corporate-owned life insurance shelter. Working paper, Arizona State University.

Chava, S., and R. Jarrow. 2004. Bankruptcy prediction with industry effects. *Review of Finance* 8 (4): 537–569.

Demers, E., and P. Joos. 2007. IPO failure risk. *Journal of Accounting Research* 45: 333–371.

Desai, M. 2003. The divergence between book income and tax income. In *Tax Policy and the Economy* (17), edited by James Poterba, 169–206. Cambridge, MA: MIT Press. Formerly "The corporate profit base, tax sheltering activity, and the changing nature of employee compensation," NBER Working Paper No. 8866.

———, and D. Dharmapala. 2006. Corporate tax avoidance and high-powered incentives. *Journal of Financial Economics* 79 (1): 145–179.

———, C. F. Foley, and J. R. Hines, Jr. 2006. The demand for tax haven operations. *Journal of Public Economics* 90 (3): 513–531.

Dyreng, S., M. Hanlon, and E. Maydew. 2008. Long-run corporate tax avoidance. *The Accounting Review* 83 (1): 61–82.

Engel, E., M. Erickson, and E. Maydew. 1999. Debt-equity hybrid securities. *Journal of Accounting Research* 37: 249–274.

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes—An Interpretation of FASB Statement No. 109*. FASB Interpretation No. 48. Norwalk, CT: FASB.

Frank, M. M., L. Lynch, and S. Rego. 2009. Tax reporting aggressiveness and its relation to aggressive financial reporting. *The Accounting Review* 84 (2): 467–496.

Frischmann, P., T. Shevlin, and R. Wilson. 2008. Economic consequences of increasing the conformity in accounting for uncertain tax benefits. *Journal of Accounting and Economics* 46: 261–278.

Gleason, C., and L. Mills. 2002. Materiality and contingent tax liability reporting. *The Accounting Review* 77 (2): 317–342.

Graham, J., and A. Tucker. 2006. Tax shelters and corporate debt policy. *Journal of Financial Economics* 81 (3): 563–594.

Graetz, M. 1999. Quoted in Tax Report. *Wall Street Journal* (February 10): A1.

Hanlon, M. 2003. What can we infer about a firm's taxable income from its financial statements? *National Tax Journal* 56 (4): 831–863.

———, and T. Shevlin. 2002. The accounting for the tax benefits of employee stock options and implications for research. *Accounting Horizons* 16 (1): 1–16.

———. 2005. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. *The Accounting Review* 80: 137–166.

———, and J. Slemrod. 2009. What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. *Journal of Public Economics* 93: 126–141.

———, and S. Heitzman. 2009. A review of tax research. Working paper, Massachusetts Institute of Technology and University of Rochester.

Hennig, C., D. Samelson, and B. Sonnier. 2008. The tax shelter dilemma: The impact of reportable transactions on the taxpayer and the tax return preparer. Working paper, Florida International University.

Hines, J. R., Jr., and E. Rice 1994. Fiscal paradise: Foreign tax havens and American business. *The Quarterly Journal of Economics* 109 (1): 149–182.

Hosmer, D., and S. Lemeshow. 2000. *Applied Logistic Regression*. 2nd edition. New York, NY: John Wiley & Sons, Inc.

Internal Revenue Service. 2005. KPMG to pay $456 million for criminal violations. Press Release IR-2005-83 (August 29): http://www.irs.gov/newsroom/article/0,,id=146999,00.html.

Jones, J. 1991. Earning management during import relief investigations. *Journal of Accounting Research* 40: 727–759.

Joos, P., and G. Plesko. 2005. Valuing loss firms. *The Accounting Review* 80 (3): 847–870.

Klein, A., and C. Marquardt. 2006. Fundamentals of accounting losses. *The Accounting Review* 81 (1): 179–206.

Kothari, S. P., A. Leone, and C. Wasley. 2005. Performance matched unexplained accrual measures. *Journal of Accounting and Economics* 39: 163–197.

Landsman, W., K. Peasnell, and C. Shakespeare. 2008. Are asset securitizations sales or loans? *The Accounting Review* 83 (5): 1251–1272.

Lisowsky, P. 2009. Inferring U.S. tax liability from financial statement information. *The Journal of the American Taxation Association* 31 (1): 29–63.

———, and W. Trautman. 2007. Book-tax consolidation, rates of return, and capital structure. *Tax Notes* (December 10): 1043–1071.

———, L. Robinson, and A. Schmidt. 2010. An examination of FIN 48: Tax shelters, auditor independence, and corporate governance. Working paper, University of Illinois at Urbana–Champaign, Tuck School of Business at Dartmouth, and Columbia University.

McGill, G., and E. Outslay. 2002. Did Enron pay taxes? Using accounting information to decipher tax status. *Tax Notes 96* (August 19): 1125–36.

———, and ———. 2004. Lost in translation: Detecting tax shelter activity in financial statements. *National Tax Journal* 57 (3): 739–756.

Mills, L. 1998. Book-tax differences and Internal Revenue Service adjustments. *Journal of Accounting Research* (Autumn): 343–356.

———, and G. Plesko. 2003. Bridging the reporting gap: A proposal for more informative reconciling of book and taxable income. *National Tax Journal* 56 (4): 865–893.

———, and K. Newberry. 2005. Firms' off-balance sheet and hybrid debt financing: Evidence from their book-tax reporting differences. *Journal of Accounting Research* 43 (2): 251–282.

———, L. Robinson, and R. Sansing. 2010. FIN 48 and tax compliance. *The Accounting Review* 85 (5): 1721-1742.

Organisation for Economic Co-Operation and Development. 2000. *Toward Global Tax Co-Operation: Report to the 2000 Ministerial Council Meeting and Recommendations by the Committee on Fiscal Affairs; Progress in Identifying and Eliminating Harmful Tax Practices*. Paris, France: OECD Publications.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22: 435–480.

Plesko, G. 2000. Evidence and theory on corporate tax shelters. National Tax Association Proceedings from the 92nd Annual Conference on Taxation, 367–371, Atlanta, GA, October.

———. 2003. An evaluation of alternative measures of corporate tax rates. *Journal of Accounting and Economics* 35: 201–226.

U.S. Department of the Treasury. 1999. *The Problem of Corporate Tax Shelters: Discussion, Analysis and Legislative Proposals. July*. Washington, D.C.: Government Printing Office.

U.S. Senate Permanent Subcommittee on Investigations, Committee on Governmental Affairs. 2003. *U.S. Tax Shelter Industry: The Role of Accountants, Lawyers, and Financial Professionals. Four KPMG Case Studies: FLIP, OPIS, BLIPS, and SC2*. November. Washington, D.C.: Government Printing Office.

Wilson, G. P. 1995. Corporate tax effectiveness: Adding value through the tax function. Working paper, Harvard University.

Wilson, R. 2009. An examination of corporate tax shelter participants. *The Accounting Review* 84 (3): 969–999.

Wysocki, P. 2004. Discussion of: Ultimate ownership, income management, and legal and extra-legal institutions. *Journal of Accounting Research* 42: 463–474.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1721–1742

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1721

FIN 48 and Tax Compliance

Lillian F. Mills
The University of Texas at Austin

Leslie A. Robinson
Dartmouth College

Richard C. Sansing
Dartmouth College and CentER, Tilburg University

ABSTRACT: We develop a model to examine the effects of Financial Accounting Standards Board (FASB) Interpretation No. 48, *Accounting for Uncertainty in Income Taxes* (FIN 48), on the strategic interaction between publicly traded corporate taxpayers and the government. Several of our findings contradict conjectures voiced by members of the business community regarding the economic effects of implementing FIN 48. Specifically, taxpayers with strong facts obtain *higher* expected payoffs from uncertain tax benefits and some disclosed liabilities *understate* the expected tax liability. Consistent with the common conjectures, however, some taxpayers are more likely to be audited or are deterred from entering into transactions that generate uncertain tax benefits because of FIN 48.

Keywords: *tax compliance; FIN 48; ASC 740-10-25; accounting for income taxes; disclosure.*

I. INTRODUCTION

Financial Accounting Standards Board (FASB) Interpretation No. 48, *Accounting for Uncertainty in Income Taxes* (FIN 48, FASB 2006), is the most significant change in financial accounting for income taxes over the past decade. This interpretation prescribes the accounting for uncertain income tax benefits recognized in a firm's financial statements in accordance with FASB Statement No. 109, *Accounting for Income Taxes*. Taxpayers are often uncertain whether the government will assess an additional income tax payment upon audit, which would reduce tax benefits claimed in the originally filed tax return. FIN 48 dictates how uncertain tax benefits should be measured and requires a taxpayer to disclose any liability for tax benefits claimed on its tax return that are not permitted to be recognized in its publicly available financial

We thank Tim Baldenius, Paul Beck, Harry Evans (editor), Corinna Ewelt, Mary Margaret Frank, Peter Frischmann, Shane Heitzman, Amy Hutton, Suzanne Paquette, George Plesko, Sonja Reiger, Terry Shevlin, our anonymous reviewers, and workshop participants at The Ohio State University, University of Oregon, Boston College, the 2008 Internal Revenue Service Research Conference, the 2008 University of North Carolina Tax Symposium, the 2008 American Taxation Association Midyear Meeting, and the 2008 American Accounting Association Annual Meeting for helpful comments.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: June 2008
Accepted: December 2009
Published Online: August 2010

statements. We investigate how the strategic interaction between publicly traded corporate taxpayers and the government changes as the result of this new accounting interpretation.

Tax uncertainty arises because of the difficulty in applying ambiguous tax laws to a set of circumstances. Shareholders and creditors previously had little information about the future cash payments corporations expected to make to the government arising from disputes over uncertain tax positions. Although the FASB intended for FIN 48 to provide financial statement users with information about tax uncertainty, the government also observes this new disclosure. Accordingly, the government receives a signal about the taxpayer's level of uncertainty about its tax filing positions. We model tax uncertainty in a strategic game between the taxpayer and government, and use a nondisclosure benchmark case to examine the effect of the FIN 48 disclosure on the strategic choices made by the taxpayer and the government.

Early models that examined the strategic interaction between the taxpayer and the government assumed that the taxpayer knew its true taxable income and that the government could discover such income with an audit. Thus, taxpayers could perfectly anticipate the consequences of an audit (Graetz et al. 1986). Beck and Jung (1989a) and Beck and Jung (1989b) examine tax compliance in an environment in which taxpayers are uncertain regarding their tax liability; the latter did so using a game theory framework in which audit probabilities are endogenous. Other studies allow the government to observe a signal regarding the taxpayer, which affects their strategic interaction (Sansing 1993; Mills and Sansing 2000; Beck et al. 2000, hereafter BDJ).[1] Our model features a *mandatory* truthful disclosure of taxpayers' private information via the new financial reporting disclosure regulation.

Specifically, we model the FIN 48 financial statement disclosure as a public signal from the taxpayer to the shareholders that also provides information to the government. As in Beck and Jung (1989b), we introduce a continuum of taxpayers that differ in the strength of their filing positions, but each taxpayer makes a discrete tax reporting decision. The government draws inferences regarding the taxpayer's type from both its tax return and its financial statements before deciding whether to audit. Our model has several features in common with BDJ. BDJ examine the effect of allowing a taxpayer to make a voluntary disclosure regarding its private information in exchange for which the government agrees to waive a penalty in case a claimed tax benefit is lost upon audit. Although the institutional setting examined in BDJ differs from the one we examine here in several important respects, our analysis of the effects of disclosure in Section IV identifies both similarities and differences in how voluntary and mandatory disclosure regimes affect taxpayer behavior and payoffs.

Our model imposes five important assumptions about the information environment. As in BDJ, we make assumptions that increase the government's ability to draw inferences from the disclosure "in order to create an environment conducive to information transfers among taxpayers and the tax agency" (BDJ, 248). Thus, our assumptions reflect our primary objective, which is to explore the *maximum* extent to which FIN 48 could affect strategic choices of both the taxpayer and the government. This approach ensures that we determine an upper bound on the effects of FIN 48; different assumptions would yield weaker effects than those we find.

First, we assume that the taxpayer files a tax return with a single government. This assumption eliminates having a single aggregate disclosure related to tax uncertainties in multiple political jurisdictions. Second, we assume that the tax return claims at most one uncertain tax benefit. Third, we assume that the government can observe whether the tax return claims an uncertain tax

[1] In BDJ, the signal is in the form of a disclosure by the taxpayer. In Mills and Sansing (2000), the signal is whether book income exceeds or equals taxable income. In Sansing (1993), the IRS privately observes a signal about the taxpayer's income, e.g., the IRS' "DIF" score.

benefit. This assumption allows the government to distinguish between taxpayers that disclose no FIN 48 liability on their financial statements because they are confident of sustaining the position and those taxpayers that did not enter into a transaction that generates an uncertain tax benefit. Fourth, we assume that the government knows the distribution from which the taxpayer's private information is drawn. This allows the government to infer the strength of the taxpayer's filing position in our model whenever the taxpayer discloses a liability associated with an uncertain tax benefit. Finally, we assume that the taxpayer reports truthfully in the sense that the firm complies fully with the financial reporting requirements of FIN 48.[2]

Our model provides important insights that are contrary to two common conjectures made by the business community regarding the effects of FIN 48. First, we show that mandatory disclosure of uncertainty about tax filing positions need not put all taxpayers at a disadvantage in the tax audit process. In fact, we show that taxpayers with strong tax positions have *higher* expected payoffs from claiming uncertain tax positions post-FIN 48. To our knowledge, the possibility that some taxpayers may benefit from the mandatory disclosure requirements of FIN 48 has not been discussed by the business community or standard-setters. Second, we show that the liability disclosed in the financial statements is not necessarily an overstatement of the expected cash payment. The conjecture that FIN 48 liabilities are likely to be overstated is due to the interpretation's departure from the traditional expected value approach.[3] Under FIN 48, the taxpayer must assume that each uncertain tax benefit is audited when calculating its tax liability. Our model illustrates that the direction of any misstatement depends on the distribution of tax benefits the taxpayer expects to retain upon audit as well as the taxpayer's expectation of the probability of being audited. For some taxpayers, however, we obtain results consistent with the common conjectures about the effect of FIN 48. Specifically, we find that some taxpayers have lower expected payoffs under FIN 48. This lower payoff arises because such taxpayers either face a higher probability of audit or refrain from entering into transactions that generate uncertain tax benefits due to the higher audit probability.

We compare the effects of mandatory disclosure pursuant to FIN 48 on taxpayer behavior and payoffs to the effects of voluntary disclosure examined in BDJ. We find that taxpayers with sufficiently strong support for the uncertain tax position do not change their behavior as a result of the mandatory disclosure and are better off because the government is less likely to audit them. Taxpayers with weaker facts are weakly better off in the voluntary disclosure setting of BDJ, but are weakly worse off in the mandatory disclosure setting that we examine. The difference between the two settings is that the government must give up something (a penalty waiver) in the voluntary setting in order to induce the taxpayer to disclose. The government does not give up anything to get the disclosure in our mandatory disclosure setting. Thus, our study enriches our understanding of the effects of disclosure on tax compliance by comparing the different effects of voluntary versus mandatory disclosure.

Not surprisingly, FIN 48 has generated broad interest and concern across the business and academic community, as well as among investors and regulators. Using our framework, concurrent

---

[2] We emphasize that our assumptions allow us to examine an important *benchmark* case. Thus, for example, while truthful reporting may not be *descriptive* per se, truthful reporting is conducive to information transfers. Not only is this important in a strategic tax compliance framework, it is also consistent with the disclosure literature. See Admati and Pfleiderer (2000, 481), who "assume, as is often done in the literature on disclosure, that all disclosed information is truthful. This is a reasonable assumption when antifraud laws are rigorously enforced. Also, disclosures are often made by third parties, such as accounting firms, whom are not directly affected by the content of the disclosure and for who the reputation for truthfulness is valuable."

[3] The expected value approach with respect to uncertain tax positions involves recording a liability equal to the probability of the uncertain tax position being audited multiplied by the expected tax deficiency, conditional on an audit occurring.

studies by Gupta et al. (2009), Dunbar et al. (2009), and Lisowsky et al. (2010) examine the effects of FIN 48 on tax reporting. We believe our model can guide current and future research during the post-FIN 48 adoption period by offering a framework for thinking about the potential economic effects of FIN 48.

In Section II, we characterize the tax and financial reporting environment in which the taxpayer operates. Section III models the strategic interaction between the taxpayer and government pre-FIN 48, which we use as our nondisclosure benchmark case. Section IV models the recognition and measurement process of FIN 48 and the strategic interaction between the taxpayer and the government post-FIN 48. In Section V, we address two common conjectures about the effects of FIN 48 in the context of our model and compare the effects of mandatory disclosure to the effects of voluntary disclosure. Section VI concludes. All proofs are in Appendix A.

II. TAX AND FINANCIAL REPORTING UNDER TAX UNCERTAINTY

Tax Reporting under Tax Uncertainty

We first characterize the corporate income tax research, filing, and settlement process to clarify the notion of tax uncertainty that we model and that FIN 48 was designed to address. The taxpayer has an opportunity to enter into a business transaction that produces an uncertain tax benefit.[4] Before deciding whether to enter into the transaction, the taxpayer privately observes all relevant facts and supporting documentation with regard to the transaction. An important aspect of the tax reporting process is that the taxpayer conducts research for purposes of applying technical tax law to the facts and circumstances of the transaction, evaluating various tax filing positions (Magro and Nutter 2009). If the taxpayer enters into the transaction, then the tax benefit is claimed on a tax return filed with the government and becomes the taxpayer's "filing position."

The government observes the tax return and thus whether the taxpayer has entered into a transaction and claimed an uncertain tax benefit, but does not observe the strength of the taxpayer's filing position. Because it is uncertain whether the taxpayer will prevail if the government challenges the claimed tax benefit, the government cannot merely assess a deficiency upon observing a claimed uncertain tax benefit; instead, it must conduct a costly audit in order to challenge the taxpayer's position. If the tax return is audited, then the outcome of a dispute could result in the taxpayer retaining all, some, or none of the tax benefit originally claimed on the tax return. Figure 1 summarizes the sequence of play.

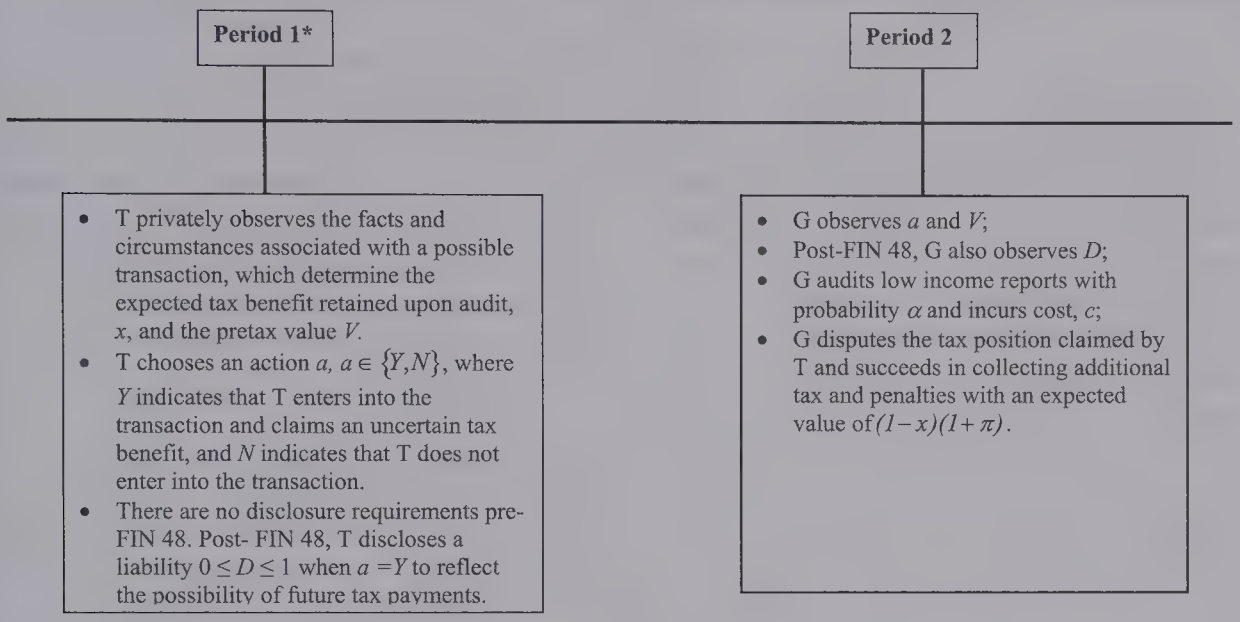
Financial Reporting under Tax Uncertainty

The financial reporting problem that FIN 48 seeks to address is how to reflect uncertain tax benefits in its financial statements during the time that potential tax disputes remain unresolved. The relevant accounting question is how much of the current tax benefit that the taxpayer receives from the uncertain tax position should be recognized as a reduction to current period income tax expense. Because we compare a FIN 48 mandatory disclosure regime to a benchmark pre-FIN 48 case of no disclosure, we first highlight important financial reporting features of the pre-FIN 48 regime.

Practice varied pre-FIN 48 with respect to recognizing liabilities for uncertain tax positions. Anecdotally, many corporations recognized a liability for uncertain tax benefits using a traditional expected value approach, although some recognized the entire tax benefit as a liability unless it was probable of being sustained. Even when the latter approach was used, the definition of

[4] Alternatively, the taxpayer may engage in a transaction that has an uncertain tax benefit for non-tax business reasons. In this case, the taxpayer's decision is not whether to engage in the transaction but is, instead, whether to claim an uncertain tax benefit associated with the transaction on the tax return. From the point of view of our model, this alternative framing has no effect.

FIGURE 1
Timeline Summarizing the Sequence of Play between the Taxpayer (T) and the Government (G)



Note that Period 1 and Period 2 do not imply two equal time periods, but rather the actions of T (Period 1) and G (Period 2). Period 1 is relatively short, while Period 2 may take many years to complete.

"probable" varied across companies. Furthermore, the extent to which corporations incorporated the probability of audit into the calculation of expected tax liabilities was unclear. Evidence in Blouin et al. (2007) suggests that firms generally recognized smaller liabilities pre-FIN 48 when they faced a lower probability of audit or detection.

Regardless of how much benefit from uncertain tax positions was recognized in financial statements pre-FIN 48, there was no explicit disclosure requirement. Thus, even if a taxpayer recognized a liability pre-FIN 48, the lack of a mandatory disclosure requirement prevented outsiders from observing the liability directly.[5] Gleason and Mills (2002) document that even large companies that have material IRS deficiencies do not voluntarily disclose the amount of their tax reserves.

FIN 48 imposes new recognition, measurement, and disclosure requirements regarding tax uncertainty. Of particular importance is that compliance with the disclosure requirement (that is, truthful reporting) makes the taxpayer's level of uncertainty transparent to the government in the post-FIN 48 regime. Furthermore, the recognition and measurement requirements serve to increase the information content of the disclosure by providing consistency in the language with which

[5] Because financial statement balance sheets only present limited detail, recorded liabilities for unrecognized tax benefits tax would not typically be visible to the public. For example, these liabilities may have been aggregated with environmental or legal liabilities. In addition, liabilities for uncertain tax positions were sometimes intermingled with deferred tax liabilities as evidenced by the prohibition in FIN 48 against this classification (FIN 48, ¶17) and the significant reclassification adjustments reported at adoption (e.g., in the first quarter of 2007, Verizon Communications Inc. reports a $3 billion reduction in deferred income taxes and a corresponding increase in other liabilities as a reclassification adjustment upon adoption of FIN 48).

taxpayers express their level of uncertainty. An important feature of the pre-FIN 48 regime was that the lack of guidance gave taxpayers no way to credibly disclose a low level of uncertainty with respect to an uncertain tax position.

III. TAX COMPLIANCE PRE-FIN 48

Pre-FIN 48 Model

Our pre-FIN 48 model examines the interaction between the taxpayer and the government prior to the implementation of FIN 48. Thus, our benchmark case is characterized by tax uncertainty without public disclosure of the taxpayer's level of uncertainty. There are two players in our model: the taxpayer (T) and the government (G). The taxpayer can take two possible actions. It can enter into a business transaction ($a = Y$) or not ($a = N$). The value of the transaction has two components, x and V, where x is the expected tax benefit associated with entering into the transaction, conditional on being audited, and V is the value of the transaction in absence of the tax benefit.

For tractability, we simplify the model so that there are only two possible tax treatments for a transaction. We normalize the best possible tax result from entering into the transaction to 1 and the worst possible tax result to 0. In order to model business transactions in which the tax benefit is decisive for the taxpayer's decision to enter into the transaction, we consider cases in which $-1 < V < 0$. Therefore, the value of the transaction when the tax-favored treatment occurs and is sustained is $V + 1 > 0$. Furthermore, the taxpayer would not enter into the transaction if it knew the tax outcome would be unfavorable because $V < 0$.[6] Consider V the pre-tax payoff *relative to a required return*; thus, we do not model transactions that lose money in raw terms, but rather, absent the tax benefit, the taxpayer would not choose this transaction. Municipal bonds and tax credit investments, for example, do not lack economic substance, but they would be negative NPV investments in absence of the tax benefits (Robinson 2010).

Thus, we consider discrete tax disputes in which rejection of the taxpayer's filing position implies a definitive alternative tax treatment. However, tax law ambiguity makes it hard for each party to determine whether the tax benefit claimed is valid. The taxpayer and the government may therefore compromise on a settlement, even though as a matter of law a court would ultimately have to allow or disallow the entire claimed tax benefit. An example is a tax-free spin-off that can only be challenged on the grounds that it lacks economic substance. A spin-off qualifies as tax-free under §355 or it does not; any decision about the applicability of tax law to the transaction must either grant or deny the claimed tax benefit in its entirety. Tax-free mergers provide a second example in which a recent tax case illustrates the discrete nature of many real world disputes. To avoid taxable gains of nearly $1.4 billion on the sale of a subsidiary, Times Mirror (the Company) engaged in complex and unconventional tax-free reverse triangular mergers with Reed Elsevier and Harcourt General in 1998. Upon audit in 2002, the Internal Revenue Service (IRS) recharacterized the reorganization as a fully taxable sale based on an argument that the transaction lacked economic substance, and assessed a tax deficiency of $550 million. The Company litigated the case and in 2006, a Tax Court judge issued a ruling that supported the IRS position disallowing tax-free treatment. As a result, the Company paid $880 million (which included interest on the deficiency) and then appealed. The IRS considered possible outcomes in the Seventh Circuit Court

[6] The notion that a taxpayer would engage in transaction only if the tax outcome was favorable can be illustrated more concretely by disclosures surrounding reorganizations. Liberty Media Group, Inc. reports the following in its 2008 annual report: "The Redemption and resulting separation of LEI from LMC are referred to as the Split Off. The Split Off is *conditioned on*, among other matters, receipt of stockholder approval and receipt of a private letter ruling from the IRS and a tax opinion from tax counsel and is expected to occur in the second quarter of 2009."

of Appeals and decided to initiate settlement. In late 2007, nearly 10 years after the transaction was entered into, the parties agreed to settle the case for $536 million (Mulligan 2007).

The taxpayer in our model is fully informed regarding the facts of the transaction, in that it knows how its particular facts fit within the applicable law, regulations, court cases, other authority, and IRS practices regarding settling appeals. Based on these facts and circumstances, the taxpayer knows the distribution of possible audit outcomes if it enters into the transaction and the IRS audits the tax return. We assume that both the taxpayer and the government are risk-neutral, which implies that the mean value (pretax value plus the expected tax benefit retained upon audit) is everything the taxpayer needs to know in order to decide whether to engage in the transaction in the pre-FIN 48 regime.[7]

Finally, we assume that no firm can make a *credible* voluntary disclosure that would deter government audits, because practice varied too much for disclosures to be meaningful until FIN 48 established consistent rules for recognition, measurement and the content of the disclosure. For example, even if a firm stated that it had recorded no liability for tax uncertainty, the government could not know whether the disclosure reflected a low probability of government detection or a strong case on its merits. Gleason and Mills (2002) find that the largest 100 nonfinancial firms rarely disclosed even the amounts of IRS disputes throughout the 1990s, even when the proposed adjustments were material. In the two quarters between FIN 48 enactment and adoption, Blouin et al. (2010) find that firms continued to not disclose contingent tax reserves. Thus, our assumption is consistent with prior empirical findings about the frequency of tax contingency disclosures prior to FIN 48.

Let x represent the expected value of the tax benefit that will be retained upon audit of a transaction that produces an uncertain tax benefit. The realization x represents the taxpayer's private information because, although the government can observe whether the taxpayer entered into a transaction that produced an uncertain tax benefit, it cannot observe the strength of the underlying position. For example, the government can tell from the tax return that the taxpayer has been a party to a merger and has treated it as tax-free, but it cannot tell how strongly the facts support this characterization. In addition, we assume that the government can observe the pretax value, V. Thus, in the pre-FIN 48 regime, taxpayers claiming the uncertain tax benefit are indistinguishable to the government in terms of their level of uncertainty. We assume that x is the realization of a random variable drawn from a uniform $[1/2, 1]$ distribution. The taxpayer observes x before deciding whether to enter into the transaction.[8] We assume $x \geq 1/2$ because we focus on cases for which the taxpayer is more likely than not to prevail in a dispute with the government. This ensures that no paid preparer has an obligation under Internal Revenue Code (IRC) §6694 to disclose any information about the transaction on the tax return.

After observing a tax return that contains an uncertain tax benefit, the government audits the tax return with probability α. The government faces an audit cost c where $c \in (0,1]$ and collects a penalty π on the expected additional tax collected $(1 - x)$, where π denotes the expected tax penalty rate $\pi \in (0, 1]$.[9] Figure 2 summarizes the expected payoffs to the taxpayer and the government when the taxpayer does and does not enter into the transaction. As illustrated in

---

[7] For purposes of our focus on tax compliance, we ignore any financial reporting preferences a taxpayer might have for either tax benefits or contingent liabilities beyond the pretax values represented in V.

[8] When we add the FIN 48 disclosure to our model, we impose additional structure on the underlying distribution of outcomes from which the FIN 48 liability is derived. We emphasize that x is the mean of the possible tax benefits retained upon audit, not the outcome itself.

[9] We impose an upper bound on the penalty rate of 100 percent to reflect the U.S. tax law, in which the penalty for negligence is 20 percent (IRC §6662) and the penalty for fraud is 75 percent (IRC §6663). Putting an upper bound on the penalty rules out implausible cases in which taxpayers with very strong positions are intimidated into filing high income reports by the presence of a draconian penalty regime.

FIGURE 2
Payoff Matrix for the Taxpayer (T) and the Government (G)

Payoff Matrix [T, G]	Audit	No Audit
Enter into transaction ($a = Y$)	$V + x - \pi(1-x), -x + \pi(1-x) - c$	$V + 1, -1$
Do not enter into transaction ($a = N$)	$0, -c$	$0, 0$

$a = Y$ when the taxpayer enters into a transaction with a pretax value, V, and an expected tax benefit retained upon audit, x. $a = N$ when a taxpayer does not enter into a transaction. π is the penalty rate, and c is the cost to the government of conducting an audit.

Figure 2, we normalize the "no transaction, no audit" payoff to 0 such that the payoff to the government from the "transaction, no audit" outcome is -1, and the difference between the government's payoff from the "transaction, no audit" outcome and the "transaction, audit" outcome is $(1 + \pi)(1 - x) - c$. Thus, the aggregate payoffs make economic sense, equaling either 0, V, $V - c$, or $-c$ in the four cells of Figure 2.

Not entering into the transaction ($a = N$) gives the taxpayer a payoff of 0. Entering into the transaction ($a = Y$) gives the taxpayer a payoff of $V + 1$ if the tax return is not audited and an expected payoff of $V + x - \pi(1 - x)$ if the tax return is audited. If the government does not conduct an audit, then the government's payoff is -1 if the taxpayer chooses $a = Y$ and 0 if the taxpayer chooses $a = N$. If it conducts an audit, then the government's payoff is $-x + \pi(1 - x) - c$ if $a = Y$ and $-c$ if $a = N$. The payoff from auditing when $a = N$ is $-c$, because audits are costly and no additional tax is collected. Thus the government never audits when $a = N$. The government's payoff from auditing the transaction, $-x + \pi(1 - x) - c$, reflects our assumption that the government may not recover all the tax on audit. In our model, audit outcomes vary because they depend on the strength of taxpayer's facts and circumstances, which determines x.

Pre-FIN 48 Equilibria

An equilibrium is an audit strategy for the government and a reporting strategy for each type-x of taxpayer, where each taxpayer's strategy is a best response to the government's strategy and the government's strategy is a best response given the set of taxpayer strategies. When the taxpayer's facts and the available authority (law, judicial precedents, and administrative guidance) make the probability of retaining more of the tax benefit relatively high, we refer to that taxpayer as having a strong position. When the taxpayer's facts and available authority make the probability of retaining more of the benefit relatively low, we refer to that taxpayer as not having a strong position.

Four types of equilibria arise. We first describe two simple equilibria in which all taxpayers claim the uncertain tax benefit. We then develop and discuss two more complex equilibria that we formalize in Proposition 1.

First, if auditing is sufficiently costly, $c \geq \frac{1+\pi}{4}$, all taxpayers enter into the transaction and the government conducts no audits. Audit costs are so high that the threat to audit is not credible; hence, all taxpayers enter into the transaction that generates uncertain tax benefits.

When audit costs are low, $c < \frac{1+\pi}{4}$, behavior depends on the pretax value V of the transaction

and in some cases the taxpayer's private information, x. Figure 3 depicts taxpayer and government behavior for the equilibria that arise when audit costs are low.

If audit costs are low and the pretax value V of the transaction is sufficiently high, $V \geq \frac{\pi-1}{2}$, then all taxpayers are willing to enter the transaction despite a certain audit. Their willingness arises because the pretax value of the transaction is sufficiently high, and the government is willing to audit all transactions with uncertain tax benefits because the audit cost is sufficiently low.

Proposition 1 below formalizes the behavior of the government and the taxpayer for the two equilibria for which the behavior of the taxpayer depends on the realization of x when $V < \frac{\pi-1}{2}$. If audit costs are low and V takes on an intermediate value, $2c - 1 < V < \frac{\pi-1}{2}$ (see Proposition 1(a)), then the government audits all tax returns that claim uncertain tax benefits and taxpayers determine their reporting strategy according to a cutoff value x^*. Taxpayers with a type-x above x^* enter into the transaction even when they are certain they will be audited, whereas taxpayers with

FIGURE 3
Pre-FIN 48 Equilibrium When $c < \frac{1+\pi}{4}$ and Post-FIN 48 Equilibrium When $x_S \leq 1 - \frac{2c}{1+\pi}$

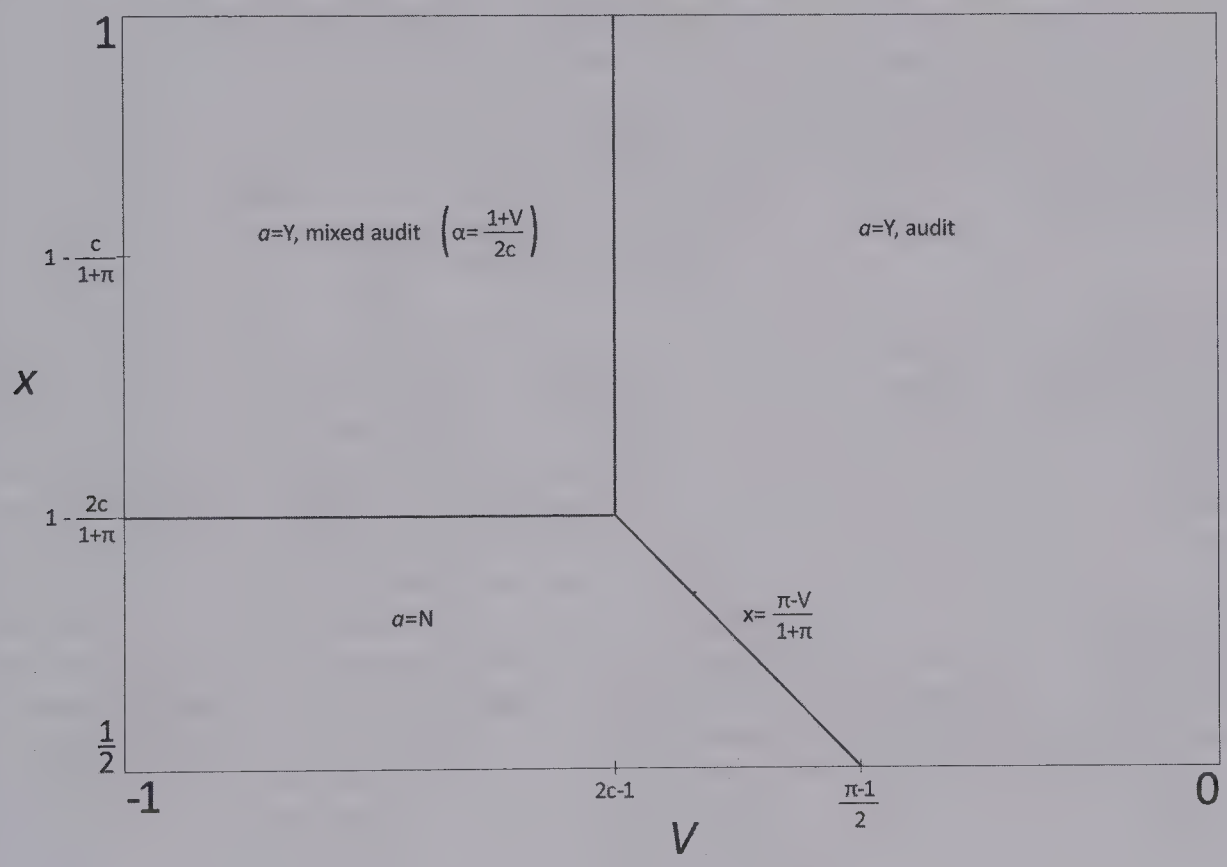


Figure 3 illustrates taxpayer and government behavior, respectively. With respect to the taxpayer, $a = Y$ when the taxpayer enters into a transaction with a pretax value, V, and an expected uncertain tax benefit retained upon audit, x. $a = N$ when a taxpayer does not enter into a transaction. With respect to the government, *audit* means the government audits all taxpayers, whereas *mixed audit* means the government audits taxpayers with probability, α. π is the penalty rate, and c is the cost to the government of conducting an audit.

a type-x below x^* do not enter into the transaction. We emphasize that taxpayers with relatively strong positions ($x > x^*$) claim an uncertain tax benefit even in the face of certain audit.

Finally, when audit costs are low and the transaction value is low, $V \leq 2c - 1$ (see Proposition 1(b)), the threat to audit deters taxpayers from entering the transaction unless $x \geq 1 - \frac{2c}{1+\pi}$. Because only taxpayers with sufficiently strong facts enter into the transaction, the government does not have an incentive to audit all tax returns that claim uncertain tax benefits. In equilibrium, the government adopts a mixed strategy, auditing some returns and not others. Given this mixed strategy, a taxpayer for whom $x = 1 - \frac{2c}{1+\pi}$ is indifferent about entering into the transaction; taxpayers with stronger facts enter into the transaction and taxpayers with weaker facts do not. Likewise, the government's expected payoff from auditing a return that claims uncertain tax benefits is equal to its payoff from not auditing the return.

Proposition 1:

(a) If audit costs (c) are low, $c < \frac{1+\pi}{4}$, and the pretax transaction value (V) is intermediate, $2c - 1 < V < \frac{\pi-1}{2}$, then:

- (i) G audits no $a = N$ tax returns,
- (ii) G audits all $a = Y$ tax returns,
- (iii) each T chooses $a = Y$ if $x > \frac{\pi-V}{1+\pi}$, and chooses $a = N$ otherwise.

(b) If $c < \frac{1+\pi}{4}$, and the pretax transaction value is low, $V \leq 2c - 1$, then:

- (i) G audits no $a = N$ tax returns,
- (ii) G audits $a = Y$ tax returns with probability $\alpha = \frac{1+V}{2c}$,
- (iii) each T chooses $a = Y$ if $x \geq 1 - \frac{2c}{1+\pi}$, and chooses $a = N$ otherwise.

IV. TAX COMPLIANCE POST-FIN 48

Post-FIN 48 Model

We next examine the interaction between the taxpayer and the government in a post-FIN 48 regime characterized by mandatory public disclosure of the taxpayer's level of uncertainty. FIN 48 determines the amount of the uncertain tax benefit the taxpayer may recognize in the financial statements under a two-step process: recognition and measurement. First, taxpayers may only recognize tax benefits that are more than 50 percent likely to be sustained by the court of last resort based solely on the technical merits of the filing position. For positions that pass the recognition step, the measurement step then determines the *amount* of the tax benefit that should be recognized, if any. The tax benefit recognized in the financial statements is the largest tax benefit that cumulatively is greater than 50 percent likely to be sustained on audit, taking into account likely settlements with the government. This two-step process is a departure from an expected value approach. Rather than recording an expected benefit (a mean) in which the expectation incorporates audit probability, both the recognition and measurement steps under FIN 48 consider a "more likely than not" threshold (a median) and assume the government has full knowledge of the position.

FIN 48 requires the taxpayer to recognize a liability for the unrecognized tax benefit to offset some or all of the total tax benefit claimed on the tax return. Because the recognized benefit is the median value of the taxpayer's retained tax benefit (the more likely than not amount), the recognized liability is equal to 1 minus the median value of the taxpayer's retained tax benefit. Most notable in our model is the requirement under FIN 48 that the taxpayer disclose the liability for the unrecognized tax benefit in its publicly available financial statements. Thus, post-FIN 48 the government can observe both an uncertain tax position in the tax return and the FIN 48 disclosure in the financial statements. This allows the government to form a more precise belief regarding the

strength of the taxpayer's filing position than it could pre-FIN 48. The government can use this revised belief to determine its audit strategy. Finally, the government can observe the value of the pretax features of the transaction, V. The cost of the audit, c, is unchanged in our post-FIN-48 model, because although the government can identify the existence of the transaction by observing a, it must still conduct a costly audit to analyze the transaction.[10]

We assume that the taxpayer's facts are sufficiently strong that they pass the recognition step. Although some transactions would not satisfy the criteria for recognition, other features of the institutional environment make this difficult to hide from the government. For example, a paid preparer has an obligation to disclose the existence of a transaction that would fail the "more likely than not" standard (IRC §6694; BDJ; Pauly 2008). Because our focus is solely on the effect of FIN 48 on the strategic interaction between publicly traded corporate taxpayers and the government, we confine our attention to settings in which the government can only condition its audit decision on the existence of the uncertain tax position on the tax return and the firm's FIN 48 disclosure.

To model the measurement processes of FIN 48, we introduce the function $m(x)$, which represents the median tax benefit retained upon audit, given x. The uncertain tax benefit recognized under FIN 48 is $1 - m(x)$. We make no assumption about whether $m(x)$ is greater or less than x; medians can be above or below means. We assume that $m(x)$ is continuous, $m'(x) \geq 0$ for all $m(x) < 1$, and $m(1) = 1$.

We allow the taxpayer to retain the full tax benefit with some positive probability $p(x)$. This implies that the expected retained benefit given some of the benefit is lost, which occurs with probability $1 - p(x)$, is $\frac{x-p(x)}{1-p(x)}$, so that the expected retained benefit is equal to x. We define x_S as the smallest value of x for which $m(x) = 1$. We denote this cutoff point as x_S to signify the point above which the taxpayer's type-x is "strong." We provide an example of a probability distribution of audit outcomes that is consistent with all of our assumptions regarding $p(x)$ and $m(x)$ in Appendix B.

We denote the disclosed FIN 48 liability as D, where $D \in [0,1]$. If $x \geq x_S$ (i.e., taxpayer has a strong position), then $m(x) = 1$ and so $D = 0$ because $D = 1 - m(x)$. If $x < x_S$ (i.e., taxpayer does not have a strong position), then $0 < D < 1$. The function $m(x)$ is common knowledge, so if $0 < D < 1$, then the government can infer x. In contrast, if $D = 0$, the government only knows that $x \geq x_S$. In summary, the mean and median settlement outcomes are linked, so we let x be the mean settlement and $m(x)$ be the median settlement. We make x the main random variable, and $m(x)$ is expressed as a function of x.

Two types of FIN 48 disclosures thus correspond with the measurement process of FIN 48: $D = 0$ or $0 < D < 1$. A disclosure of $D = 0$ in the financial statements is consistent with two situations. First, $D = 0$ when $a = N$, because the taxpayer did not claim an uncertain tax benefit. Second, $D = 0$ when the taxpayer claims an uncertain tax benefit and has a strong position (i.e., the taxpayer records $D = 1 - m(x)$ where $m(x) = 1$). Because the government can observe from

[10] Recall that we assume there is only one transaction in our model, so a FIN 48 disclosure fully reveals which transaction is uncertain. For most companies of any complexity, the FIN 48 disclosure will be an aggregate liability for multiple tax positions in multiple taxing jurisdictions. Although FASB's initial exposure draft proposed requiring corporations to disclose domestic and foreign components of unrecognized tax benefits separately, that finer disclosure did not survive the adopted standard. Thus, even if the FIN 48 disclosure helps the IRS select the taxpayer for audit, the FIN 48 disclosure does not lower the cost of conducting the audit. An important real world example that supports this feature in our model is the recent Textron case (*Textron v. Comm.* 2007), in which the court disallowed the IRS access to workpapers prepared during an audit. The business community has speculated that one of the key ways that the government will reduce their audit cost as a result of FIN 48 is by obtaining documentation used to prepare the FIN 48 financial statement disclosure. The Textron case seems to provide taxpayers with support for asserting that the work product privilege (*Arthur Young v. Comm.* 1984) applies to tax accrual workpapers that were provided to the taxpayer's independent auditors.

the tax return whether the taxpayer has claimed an uncertain tax benefit, it can distinguish between these two situations for which $D = 0$.

We assume the independent audit process ensures that the disclosure D is truthful, i.e., consistent with FIN 48. As a result, taxpayers without strong positions cannot simply decide to disclose $D = 0$ to pool with taxpayers with strong positions. Introducing another actor to the game (the independent auditor) would complicate our model and obscure our goal for this initial study. Our goal from imposing this simplifying assumption is to strengthen the government's information set so that we can evaluate the maximum effect of FIN 48 on the strategic interaction between the taxpayer and the government.

We depict the possible realizations of x and associated value of D in Figure 4.

Post-FIN 48 Equilibria

We consider the equilibria that can occur in two cases: (1) $x \geq x_S$, and (2) $x < x_S$. When $x \geq x_S$, the taxpayer has a strong position, and we begin our analysis with this case.

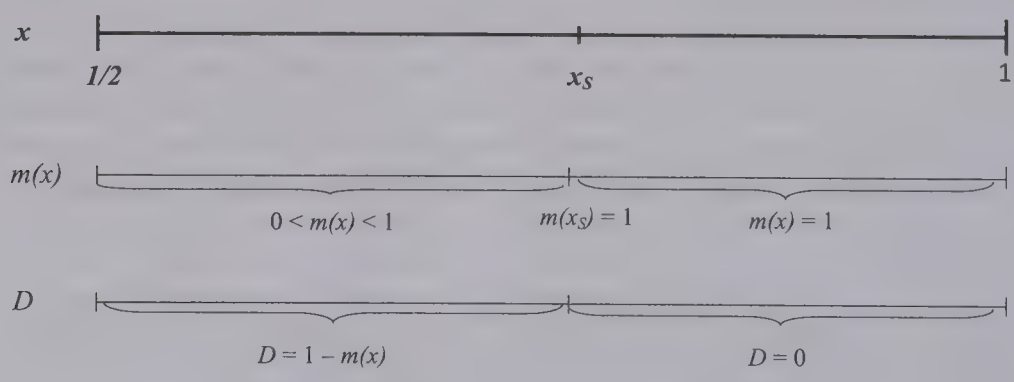
Taxpayer with a Strong Position $(x \geq x_S)$

A taxpayer that claims an uncertain tax benefit with a strong position discloses $D = 0$. When the government observes $a = Y$ and $D = 0$, it knows that x lies in the interval $[x_S, 1]$. As before, we first describe two simple equilibria in which taxpayers always claim the uncertain tax position. We formalize more complex cases for taxpayers with strong positions in Proposition 2.

If x_S is sufficiently high, $x_S \geq 1 - \frac{2c}{1+\pi}$, then the disclosure $D = 0$ deters the government from auditing. With no government audits, all taxpayers with strong positions claim the tax benefit.

If $x_S < 1 - \frac{2c}{1+\pi}$, then the strategies of the taxpayer and the government depend on the pretax value V of the transaction and the realization of x. When the value of the transaction is high, $V \geq \frac{\pi-1}{2}$, a taxpayer with a strong position is willing to claim the tax benefit even if an audit is certain. In this case, all taxpayers with a strong position claim the tax benefit, and the government audits all returns that claim the tax benefit.

FIGURE 4
Possible Realizations of the Expected Tax Benefit, x, and Associated Values of the Disclosed FIN 48 Liability, D



The expected tax benefit retained on audit is x. The median retained benefit on audit is $m(x)$. $D = 0$ is disclosed when taxpayers have strong positions, which implies $x \geq x_S$ in our model. $0 < D < 1$, or $1 - m(x)$, is disclosed when taxpayers do not have strong positions, which implies that $\frac{1}{2} < x < x_S$ in our model.

When the value of the transaction is intermediate, $2c - 1 < V < \frac{\pi-1}{2}$, or low, $V \leq 2c - 1$, the behavior of the taxpayer depends on the realization of x. We characterize the behavior of the taxpayer with a strong position and the government in Proposition 2 below.

When V takes on an intermediate value, $2c - 1 < V < \frac{\pi-1}{2}$, (see Proposition 2(a)), the government audits all tax returns that claim uncertain tax benefits and taxpayers determine their reporting strategy according to a cutoff value x^*. Taxpayers with a type-x above x^* enter into the transaction even when they are certain they will be audited, while taxpayers with a type-x below x^* do not enter into the transaction.

When the transaction value is low, $V \leq 2c - 1$ (see Proposition 2(b)), the threat to audit deters taxpayers from entering the transaction unless $x \geq 1 - \frac{2c}{1+\pi}$. Because only taxpayers with sufficiently strong facts enter into the transaction, the government does not have an incentive to audit all tax returns that claim uncertain tax benefits. In equilibrium, the government adopts a mixed strategy, auditing some returns and not others.

Proposition 2: When $x \geq x_S$, $D = 0$ for both $a = Y$ and $a = N$.

(a) If the lowest value of x for which no liability is recognized under FIN 48 is low $x_S < 1 - \frac{2c}{1+\pi}$, and the value ($V$) of the transaction is intermediate, $2c - 1 < V < \frac{\pi-1}{2}$, then:
  (i) G audits no $a = N$ tax returns,
  (ii) G audits all $a = Y$ tax returns, and
  (iii) each T chooses $a = Y$ if $x > \frac{\pi-V}{1+\pi}$ and chooses $a = N$ otherwise.

(b) If $x_S < 1 - \frac{2c}{1+\pi}$ and the value of the transaction is low, $V \leq 2c - 1$, then:
  (i) G audits no $a = N$ tax returns,
  (ii) G audits $a = Y$ tax returns with probability $\alpha = \frac{1+V}{2c}$, and
  (iii) each T chooses $a = Y$ if $x > 1 - \frac{2c}{1+\pi}$ and chooses $a = N$ otherwise.

Taxpayer without a Strong Position $(x < x_S)$

When $x < x_S$, the taxpayer is more likely than not to lose part of the claimed tax benefit. When a taxpayer without a strong position enters into a transaction that generates an uncertain tax benefit, it must disclose $0 < D < 1$, where $D = 1 - m(x)$. Unlike the case in which $D = 0$, when the government observes $a = Y$ and $0 < D < 1$, the taxpayer's private information, x, is fully revealed to the government because the government knows $m(x)$. In this setting, both the taxpayer and the government are in a full information environment. We characterize the behavior of the taxpayer without a strong position and the government in Proposition 3 below.

Proposition 3: When $\frac{1}{2} < x < x_S$, $D = 1 - m(x)$ when $a = Y$ and $D = 0$ when $a = N$.

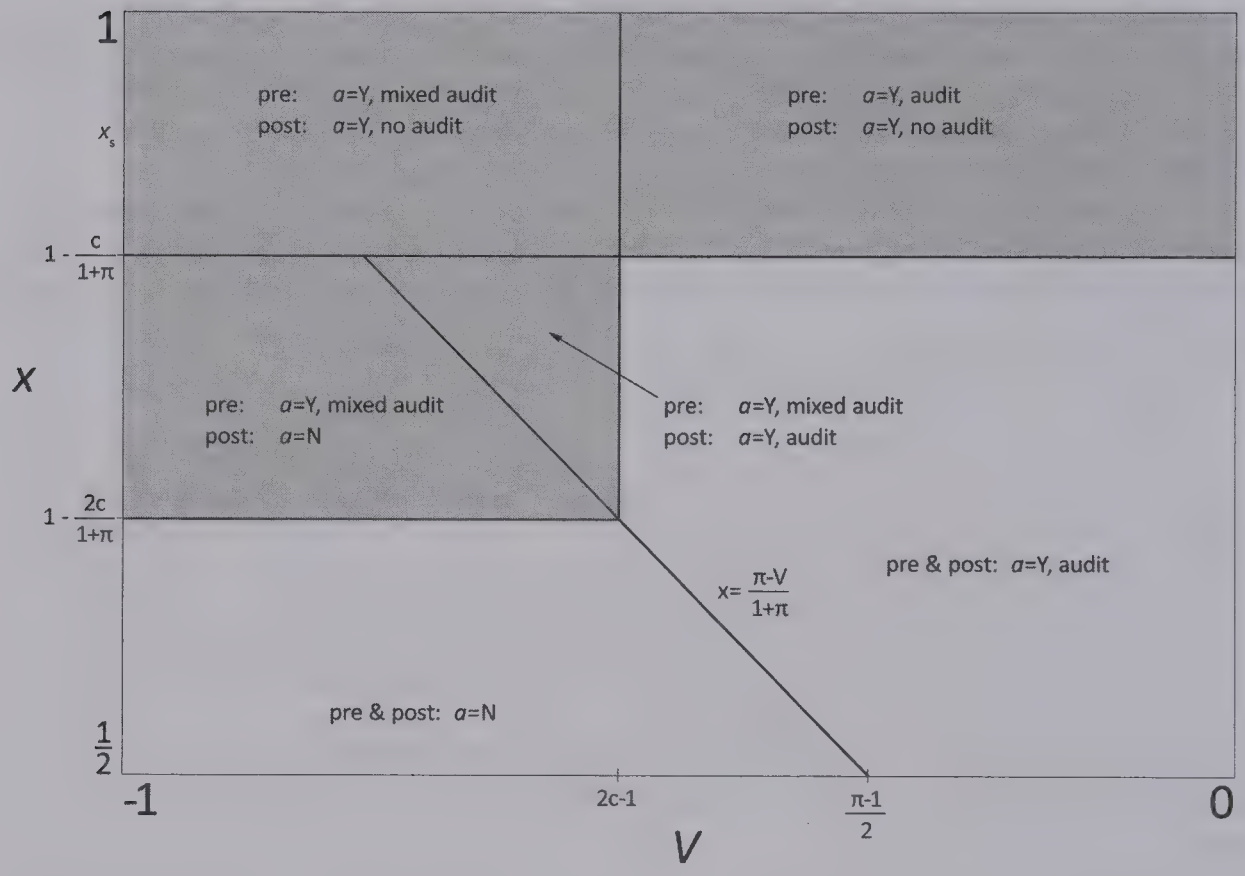
(a) If $x \geq 1 - \frac{c}{1+\pi}$ then:
  (i) G audits no tax returns, and
  (ii) all T chooses $a = Y$.

(b) If $\frac{1}{2} \leq x < 1 - \frac{c}{1+\pi}$, then:
  (i) G audits no $a = N$ tax returns,
  (ii) G audits all $a = Y$ tax returns, and
  (iii) each T chooses $a = Y$ if $x > \frac{\pi-V}{1+\pi}$ and chooses $a = N$ otherwise.

We illustrate the behavior of taxpayers and the government when $c < \frac{1+\pi}{2}$ in Figures 3, 5, and 6. Figure 3 represents the situation in which $x_S \leq 1 - \frac{2c}{1+\pi}$. Recall also that Figure 3 depicts the pre-FIN 48 equilibria. Figure 5 represents $x_S \geq 1 - \frac{c}{1+\pi}$, and Figure 6 represents $1 - \frac{2c}{1+\pi} < x_S < 1 - \frac{c}{1+\pi}$.

We emphasize that the region of each figure for which $x \geq x_S$ is the region for which all taxpayers that engage in the transaction disclose $D = 0$ and the region for which $x < x_S$ is the region for which all taxpayers that engage in the transaction disclose $D = 1 - m(x) > 0$. Thus,

FIGURE 5
Comparing Equilibria Pre-FIN 48 and Post-FIN 48 When $x_S \geq 1 - \frac{c}{1+\pi}$

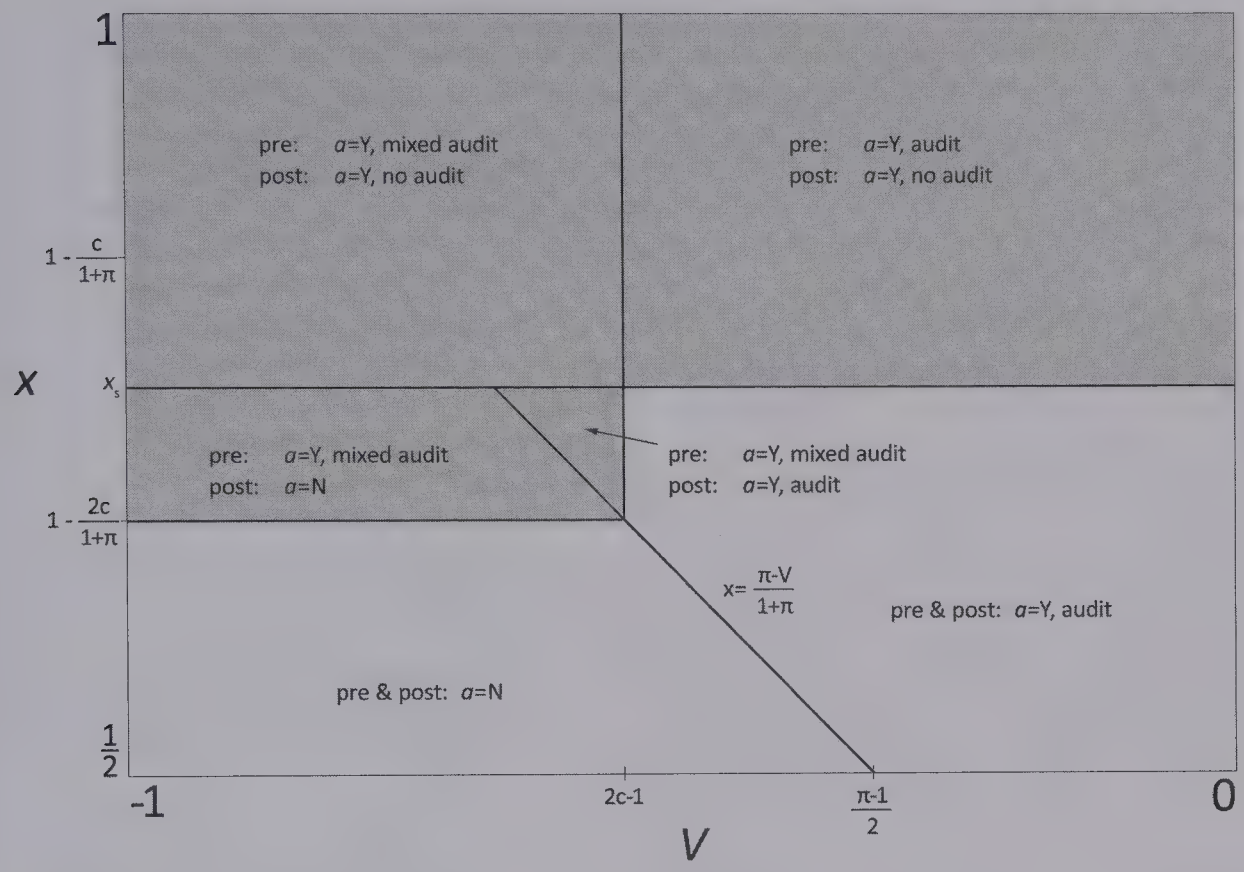


because Figure 5 shows a high range for x_S, it depicts the situation in which audit costs are relatively high, whereas Figure 3 represents a situation in which audit costs are relatively low.

Figure 3 characterizes the behavior of taxpayers and the government both prior to FIN 48 and after FIN 48 when $x_S \leq 1 - \frac{2c}{1+\pi}$. Most taxpayers disclose no FIN 48 liability because x_S is so low. In the extreme case in which $x_S = \frac{1}{2}$, the disclosure has no information content because all taxpayers make the same disclosure. For taxpayers that disclose a FIN 48 liability, the audit costs are so low that the government does not need the disclosure to convince it to audit. Therefore, the FIN 48 disclosure does not affect either taxpayer or government behavior when $x_S \leq 1 - \frac{2c}{1+\pi}$.

Figures 5 and 6 illustrate the behavior of the government and the taxpayers both pre-FIN 48 and post-FIN 48. The figures are similar, in that each features a threshold value, x^*, for which post-FIN 48 all $x \geq x^*$ enter into the transaction and the government does not audit. For all $x < x^*$, the taxpayer enters into the transaction (and the government audits) post-FIN 48 when V is high and does not enter into the transaction when V is low. The only difference between Figures 5 and 6 is that the threshold value x^* is less than x_S in Figure 5 but is equal to x_S in Figure 6. This difference is driven by audit costs. In Figure 5, some taxpayers enter into the transaction despite disclosing a FIN 48 liability. In Figure 6, audit costs are low enough that taxpayers are deterred from entering into the transaction unless they have no FIN 48 liability. We interpret these figures when discussing our insights in Section V.

FIGURE 6
Comparing Equilibria Pre-FIN 48 and Post-FIN 48 When $1-\frac{2c}{1+\pi} < x_S < 1-\frac{c}{1+\pi}$



V. INSIGHTS AND CONTRIBUTIONS

Our model contributes to our understanding of FIN 48 by challenging commonly expressed beliefs in the business press regarding the effects of FIN 48. It also highlights the differences between the effects of mandatory and voluntary disclosure regimes. First, we show how some of our results contradict conjectures by the business press. Second, we compare our results to those from the BDJ model of voluntary disclosure in a strategic tax compliance setting.

Challenging Common FIN 48 Conjectures

A common conjecture about the effects of FIN 48 is that the new accounting interpretation will typically harm taxpayers and aid the government by providing a "road map" to a taxpayer's controversial tax return filing positions. In discussing FIN 48 during an SEC speech, Chester Spatt, Chief Economist and Director of the Office of Economic Analysis of the SEC, stated that "providing publicly more information about the taxpayer's position on salient tax issues may provide a 'roadmap' for the government that may undercut the firm's bargaining power in the associated tax disputes" (Spatt 2007).[11] Similarly, Allergan, Inc.'s comment letter to the FASB

[11] Available online at: http://www.sec.gov/news/speech/2007/spch030807css.htm as of July 22, 2010, Speech by SEC staff: "The Economics of FIN 48: Accounting for Uncertainty in Income Taxes."

argues that disclosures of accrued liability "will provide a readily visible red flag to all taxing authorities as to the magnitude of potential audit issues for which the taxing authorities should be looking."[12]

By comparing the behavior of taxpayers and the government pre- and post-FIN 48, we show that some taxpayers that claim uncertain tax benefits with relatively strong positions are better off post-FIN 48. The shaded regions in Figures 5 and 6 are those in which the equilibrium behavior changes due to FIN 48. Consider Figure 5. When $x \geq 1 - \frac{c}{1+\pi}$ and the FIN 48 disclosure reveals that x is high, the "road map" makes taxpayers better off because the government does not audit the transaction and makes the government better off because it does not audit taxpayers with strong positions. FIN 48 provides a credible signal that the taxpayer has relatively strong facts—a signal that makes both parties better off. The effects of FIN 48 are more subtle when $1 - \frac{2c}{1+\pi} < x < 1 - \frac{c}{1+\pi}$ and $V < 2c - 1$ If $x \leq \frac{\pi - V}{1+\pi}$, then the FIN 48 disclosure deters the taxpayer from engaging in the transaction. This makes the government better off and the taxpayer worse off. On the other hand, if $x > \frac{\pi - V}{1+\pi}$, then the taxpayer is worse off because it faces certain audit, but the government is better off because the government conditions its audit on the FIN 48 disclosure. Figure 6 is similar to Figure 5; the only difference is the size of the regions. Specifically, the size of the shaded areas in which both parties are better off increases and the size of the shaded areas in which the government is better off but the taxpayer is either indifferent or worse off decreases.

In contrast, Figure 3 reveals that the behavior of neither the taxpayer nor the government changes when audit costs are sufficiently low. When $x_S < 1 - \frac{2c}{1+\pi}$, most taxpayers disclose $D = 0$, either because they do not engage in the transaction or because $x \geq x_S$. Because most taxpayers make the same FIN 48 disclosure, it is of little use as a signal. Conversely, taxpayers with $D > 0$ were being audited in the pre-FIN 48 regime, so the disclosure of their relatively weak facts has no effect on their probability of being audited.

Finally, we consider the effects of FIN 48 when $c \geq \frac{1+\pi}{4}$. There are two possibilities, which we explain but do not illustrate in a figure. If $c \geq \frac{1+\pi}{2}$, then auditing is too costly in either the pre-FIN 48 or post-FIN 48 environments. In each case, every taxpayer engages in the transaction and the government does not audit. If $\frac{1+\pi}{4} \leq c < \frac{1+\pi}{2}$, then auditing is too costly pre-FIN 48 but not necessarily post-FIN 48 because the government can condition its audit decision on the taxpayer's FIN 48 disclosure. In absence of a FIN 48 disclosure, the government lacks a credible threat to audit any taxpayer if it would be unwilling to audit the average taxpayer $\left(x = \frac{3}{4}\right)$. With a FIN 48 disclosure, the government is weakly better off because it audits some taxpayers that do not have strong facts and deters other taxpayers from claiming the uncertain tax position. Taxpayers are weakly worse off.

Overall, we see that FIN 48 can make the government better off but never makes it worse off. The FIN 48 "road map" is weakly valuable to the government. But the road map can also be valuable to the taxpayer, deterring audits with low expected payoffs to the government. Of course, the road map can also make the taxpayer worse off, either by deterring the taxpayer from entering into a transaction with uncertain tax benefits or by subjecting the taxpayer to a higher audit probability.

A second common conjecture is that FIN 48 may cause many issuers of GAAP financial statements to significantly overstate their ultimate tax liabilities (BNA 2006). Critics of FIN 48 note that because the recognition and measurement process of FIN 48 assumes that the government will audit a particular tax return filing position, the departure from expected value will cause the recorded liability to be greater than the expected tax payment. J. Howard Stecker, the V.P. and

[12] Available online at: http://www.fasb.org/jsp/FASB/CommentLetter_C/CommentLetterPage?project_id=1215-001 as of July 22, 2010.

Chief Tax Officer of Prudential Financial, wrote a comment letter to the FASB stating that the proposed interpretation "will have the effect of consistently overstating the tax accruals for uncertain tax positions [relative to the responsible professional judgment under FAS 5] in direct conflict with the conceptual framework of the Board."[13]

Recall that x is the expected value of the tax benefit associated with an uncertain tax benefit that will be retained upon audit, α is the probability of being audited, and D is the difference between the "as filed" tax benefit of 1 and the median tax benefit retained on audit. Therefore, the expected tax liability is $\alpha(1 - x)$; however, the disclosed FIN 48 liability, D, is either 1, 0, or $1 - m(x)$. We denote a FIN 48 liability as either *overstated* when the disclosed liability exceeds the expected liability, *understated* when the expected liability exceeds the disclosed liability, or as *correctly* stated when the two are equal.

When $D = 0$, the disclosed liability is correctly stated when $\alpha = 0$ and is understated otherwise. When $D = 1$, the disclosed liability is overstated. There are two factors that determine whether the disclosed liability is overstated or understated when $D = 1 - m(x)$: the distribution of expected tax benefit retained upon audit and the probability of audit. We explore both factors in turn. First, the implications of the distribution of the expected tax benefit retained on audit can be examined by comparing the expected liability, $1 - x$, to the disclosed liability, $1 - m(x)$, conditional upon audit (i.e., when $\alpha = 1$). These two liabilities can differ because the median retained tax benefit, $m(x)$, can be higher or lower than the mean retained tax benefit, x. Second, the implications of the probability of an audit can be addressed by comparing the expected liability, $\alpha(1 - x)$ to the disclosed liability, $1 - m(x)$. The expected liability conditional upon audit, $1 - x$, decreases when $\alpha < 1$. However, $\alpha < 1$ does not change the amount of the disclosed liability because FIN 48 does not incorporate audit probabilities into the measurement process. Thus, when the disclosed liability, $1 - m(x)$, exceeds the expected liability conditional upon audit, $1 - x$, $\alpha < 1$ exacerbates the amount by which the disclosed liability is overstated. However, when the expected liability conditional upon audit, $1 - x$, exceeds the disclosed liability, $1 - m(x)$, $\alpha < 1$ results in a disclosed liability that could be higher, lower, or equal to the expected liability, depending on the equilibrium value of α.

Comparison to Voluntary Disclosure Research

BDJ analyze the effect of voluntary taxpayer disclosure on the strategic interaction between the taxpayer and the government. Specifically, BDJ model the effect of allowing the government to waive a penalty if the taxpayer discloses the existence of a transaction that generates uncertain tax benefits. An important objective of BDJ is to determine the conditions under which taxpayer disclosure would occur. The disclosure in BDJ, if made, truthfully reveals the existence of the transaction rather than the strength of the taxpayer's position; thus, in equilibrium, the decision to disclose provides only a coarse signal regarding the taxpayer's private information. BDJ show that the disclosure can benefit both sides of the transaction because the taxpayer can avoid penalties, whereas the government enjoys lower audit costs.

In contrast, we model mandatory disclosure such that examining the conditions under which disclosure occurs is not relevant. An important difference of mandatory disclosure is that the disclosure itself does not *directly* affect the payoffs associated with the players' possible actions; recall voluntary disclosures are made in BDJ to obtain penalty reductions. Furthermore, the mandatory disclosure pursuant to FIN 48 allows the government to learn the strength of the taxpayer's

[13] Available online at: http://www.fasb.org/jsp/FASB/CommentLetter_C/CommentLetterPage?project_id=1215-001 as of July 22, 2010.

uncertain tax position directly, instead of inferring it from the equilibrium disclosure decisions of all taxpayers.

Although the voluntary and mandatory disclosure settings feature important institutional differences, the effects of disclosure in each setting have both similarities and differences. Taxpayers with sufficiently strong facts (low p in BDJ, high x in our study) do not change their behavior when disclosure is introduced in either the voluntary or mandatory disclosure settings, but actually benefit from the possibility of disclosure in each setting because they are audited less than if no disclosure was possible. In BDJ, taxpayers with strong facts claim the uncertain tax benefit and do not disclose the transaction (i.e., the expected value of the penalty reduction is so low given the strength of their position that it does not outweigh the possibility of providing a roadmap to the government). However, the fact that other taxpayers do disclose induces the government to audit taxpayers with strong facts less frequently. In the mandatory disclosure setting, it is the direct disclosure of the strong taxpayers' facts that deters the government from auditing such taxpayers.

Some taxpayers with weaker facts do change their behavior when disclosure is introduced, but in quite different ways in the two settings. When disclosure is voluntary, the possibility of disclosure induces some of these taxpayers to claim the tax benefit and disclose the transaction; the possibility of disclosure makes these taxpayers better off. In contrast, taxpayers with weaker facts are weakly worse off in the mandatory disclosure setting, as their disclosure induces weakly more aggressive auditing from the government. The difference in the two settings is that the government must give up something (a penalty waiver) in the voluntary setting in order to induce the taxpayer to disclose. The government does not give up anything to get the disclosure in the mandatory disclosure setting. Thus, taken together, BDJ and our study create a cohesive understanding of the effects of voluntary and mandatory disclosure on strategic tax compliance.

VI. CONCLUSION

We investigate how the strategic interaction between publicly traded corporate taxpayers and the government changes as a result of a new accounting interpretation, FIN 48, which requires these taxpayers to disclose liabilities in public financial statements for uncertain tax benefits claimed on a tax return. Our model provides two interesting insights that could not be gleaned from a cursory examination of FIN 48.

First, we show that taxpayers are not necessarily harmed by FIN 48. In fact, we find that taxpayers with strong positions have higher expected payoffs post-FIN 48. Although the disclosed liability does provide a roadmap in our simple one-position, one-jurisdiction model, mandatory disclosure permits taxpayers with strong facts to benefit through truthful reporting of little or no liability. This positive effect of FIN 48 for some taxpayers has not been documented to date. In addition, taxpayers with weaker facts are not necessarily harmed by FIN 48, even when the disclosure provides a "perfect roadmap" for the government. This observation results from the nuanced relation between x and c, whereby uncertain tax benefits claimed cannot be extracted from taxpayers except through a costly audit, settlement, and collection process.

Second, we find that the liability disclosed under FIN 48 can be overstated or understated relative to the expected cash payment. In contrast, International Financial Reporting Standards (IFRS) differ from U.S. Generally Accepted Accounting Principles (GAAP) by using an expected value approach, which differs from the two-step recognition and measurement process under GAAP. If expected values are more relevant to shareholders, then convergence toward IFRS standards, accompanied by disclosure, could potentially improve financial reporting.

Our model imposes certain simplifying assumptions. Critical among these is truthful reporting of the FIN 48 liability, because we assume that independent auditors and regulators assure such reporting in the post-Sarbanes-Oxley period. If independent auditors assure such truthful reporting and if the IRS relies upon it, then one effect of FIN 48 appears to be the substitution of indepen-

dent auditors for the government. That is, if the independent auditors decide that the tax position is so strong that that all the benefit can be recognized and no liability recorded, then the IRS can forgo examination. Because all public corporations receive independent audits of their GAAP financial statements but the IRS does not have the budget to examine all taxpayers, shifting this role could increase overall tax compliance by directing IRS attention to taxpayers with the weakest cases and deterring some aggressive tax claims. To the extent that mandatory disclosure and post-SOX auditing rigor requires independent auditors to evaluate the merits of uncertain tax benefits, government audit costs could decrease or government audits will be targeted more effectively. However, there may be a higher deadweight social cost due to a higher independent audit costs. Future analytical work could consider incentives and constraints for taxpayers to falsify their disclosures.

We believe our model also has various testable empirical implications. Over time, researchers could study whether the IRS more frequently audits returns with high disclosed liabilities relative to claimed tax benefits or whether FIN 48 reduces the amount of tax planning done in large public companies. In the short run, experiments and field studies could contribute evidence to test some of the implications of our model.

APPENDIX A

Proof of Proposition 1 (Pre-FIN 48)

(a)(i) Not auditing when $a = N$ is a dominant strategy for the government because $c > 0$.

(a)(ii) The government will audit when $a = Y$ if $E[-x + \pi(1 - x) - c] > -1$. Given that only taxpayers with $x \geq \frac{\pi - V}{1+\pi}$ choose $a = Y$, $E[x \mid a = Y] = \frac{1+2\pi-V}{2+2\pi}$. Therefore, the government's expected payoff from auditing when $a = Y$ is $\frac{V-1}{2} - c \geq -1$ because $V > 2c - 1$. Therefore, the government audits all tax returns of $a = Y$.

(a)(iii) A taxpayer of type x will choose $a = Y$ if and only if the expected payoff is greater than 0, the payoff from choosing $a = N$. Given that the government audits all tax returns of $a = Y$, the taxpayer's expected payoff from reporting $a = Y$ is $E[V + x - \pi(1 - x)] \geq 0$ for all $x \geq \frac{\pi - V}{1+\pi}$. Therefore, taxpayers for whom $x \geq \frac{\pi - V}{1+\pi}$ choose $a = Y$, and other taxpayers choose $a = N$.

(b)(i) Same argument as in (a)(i).

(b)(ii) Given that only taxpayers with $x \geq 1 - \frac{2c}{1+\pi}$ choose $a = Y$, $E[x \mid a = Y] = \frac{1-c+\pi}{1+\pi}$. The government is indifferent between auditing and not auditing tax returns of $a = Y$ because $E[-x + \pi(1 - x) - c] = -1$. Therefore, the government's strategy of auditing tax returns of $a = Y$ with probability $\alpha = \frac{1+V}{2c}$ is a best response. Because $V < 2c - 1$, α is bounded above by 1 and because $V > -1$, α is bounded below by 0.

(b)(iii) A taxpayer of type x will choose $a = Y$ if and only if the expected payoff is greater than 0, the payoff from choosing $a = N$. Given that the government audits when $a = Y$ with probability $\alpha = \frac{1+V}{2c}$, the expected payoff from choosing $a = Y$ is $\alpha[V + x - \pi(1 - x)] + (1 - \alpha)[V + 1] = \frac{(1+V)[2c-(1+\pi)(1-x)]}{2c} \geq 0$ for all $x \geq 1 - \frac{2c}{1+\pi}$. Because $c < \frac{1+\pi}{4}$ the cutoff-value for x is greater than $\frac{1}{2}$. QED

Proof of Proposition 2 (Post-FIN 48: $x \geq x_S$)

(a)(i) Not auditing when $a = N$ is a dominant strategy for the government because $c > 0$.

(a)(ii) The government will audit when $a = Y$ if $E[-x + \pi(1 - x) - c] > -1$. Given all taxpayers for whom $x \geq \frac{\pi - V}{1+\pi}$ choose $a = Y$, the government's expected payoff from auditing when $a = Y$ is $\frac{V-1-2c}{2} \geq -1$ because $V \geq 2c - 1$, and $E[x \mid x \geq \frac{\pi - V}{1+\pi}] = \frac{1+2\pi-V}{2(1+\pi)}$. Therefore, the government audits whenever $a = Y$.

(a)(iii) A taxpayer of type x will choose $a = Y$ if and only if the expected payoff is greater than 0, the payoff from choosing $a = N$. Given that the government audits all tax returns of $a = Y$, the expected payoff to the taxpayer when $a = Y$ is $E[V + x - \pi(1 - x)] > 0$ for all $x \geq \frac{\pi - V}{1+\pi}$.

(b)(i) Same argument as in (a)(i).

(b)(ii) Given that only taxpayers for whom $x \geq 1 - \frac{2c}{1+\pi}$ choose $a = Y$, $E[x \mid a = Y] = \frac{1-c+\pi}{1+\pi}$. The government is indifferent between auditing and not auditing tax returns of $a = Y$ because $E[-x + \pi(1 - x) - c] = -1$. Therefore, the government's strategy of auditing tax returns of $a = Y$ with probability $\alpha = \frac{1+V}{2c}$ is a best response. Because $V < 2c - 1$, α is bounded above by 1 and because $V > -1$, α is bounded below by 0.

(b)(iii) A taxpayer of type x will choose $a = Y$ if and only if the expected payoff is greater than 0, the payoff from choosing $a = N$. Given that the government audits tax returns of $a = Y$ with probability $\alpha = \frac{1+V}{2c}$, the expected payoff from reporting $a = Y$ is $\alpha[V + x - \pi(1 - x)] + (1 - \alpha)[V + 1] = \frac{(1+V)[2c-(1+\pi)(1-x)]}{2c} \geq 0$ for all $x \geq 1 - \frac{2c}{1+\pi}$. The cutoff-value for x is greater than x_S because $x_S < 1 - \frac{2c}{1+\pi}$. QED

Proof of Proposition 3 (Post-FIN 48: $\frac{1}{2} < x < x_S$)

(a)(i) Not auditing when $a = N$ is a dominant strategy for the government because $c > 0$.

(a)(ii) The government will audit when $a = Y$ if $E[-x + \pi(1 - x) - c] > -1$. Auditing all tax returns of $a = Y$ is a best response by the government because $x < \frac{1+\pi-c}{1+\pi}$.

(a)(iii) A taxpayer of type-x will choose $a = Y$ if and only if the expected payoff is greater than 0, the payoff from choosing $a = N$. Given that the government audits all tax returns of $a = Y$, the expected payoff from choosing $a = Y$ is $V + x - \pi(1 - x) \geq 0$, which is positive if and only if $x \geq \frac{\pi - V}{1+\pi}$.

(b)(i) Same argument as in (a)(i). The government will audit a tax return of $a = Y$ if $\pi(1 - x) - x - c \geq -1$. Not auditing when $a = Y$ is a best response by the government because $x \geq \frac{1+\pi-c}{1+\pi}$.

(b)(ii) Given that the government does not audit any tax returns, all taxpayers prefer $a = Y$ to $a = N$. QED

APPENDIX B

To model the measurement processes of FIN 48, we introduce the function $m(x)$, which represents the median tax benefit retained upon audit, given x. The uncertain tax benefit recognized under FIN 48 is equal to $1 - m(x)$. We make no assumption about whether $m(x)$ is greater or less than x; medians can be above or below means. We assume that $m(x)$ is continuous, $m(0) = 0$, and $m(1) = 1$. We also assume that $m'(x) \geq 0$ for all $m(x) < 1$. Below, we provide an example of a probability distribution of audit outcomes that is consistent with all of these assumptions.

Suppose a firm with private information x retains a tax benefit y that is equal to 1 with probability $p(x) = x^2$, and drawn from a distribution $f(y) = x[y^{x-1}]$, $0 < y < 1$ otherwise. This distribution has the following properties. First:

$$\int_0^1 x[y^{x-1}]dy = 1,$$

which is required for a valid probability density function. Second:

$$p(x) + [1 - p(x)] \int_0^1 xy[y^{x-1}]dy = x,$$

so the expected retained tax benefit upon audit is equal to the value of the taxpayer's private information. Third, the median of the distribution, $m(x)$, is the value of y that solves:

$$[1 - p(x)] \int_0^{m(x)} x[y^{x-1}]dy = \frac{1}{2},$$

which has a solution of $m(x) = [2(1 - x^2)]^{-1/x}$, $1/2 \leq x < \sqrt{2}/2$, and $m(x) = 1$ if $\sqrt{2}/2 \leq x \leq 1$, which has all the assumed properties of the $m(x)$ function.

REFERENCES

Admati, A., and P. Pfleiderer. 2000. Forcing firms to talk: Financial disclosure regulation and externalities. *Review of Financial Studies* 13 (3): 479–519.

Beck, P., and W. Jung. 1989a. Taxpayer compliance under uncertainty. *Journal of Accounting and Public Policy* 8 (Spring): 1–27.

———, and ———. 1989b. Taxpayers reporting decisions and auditing under information asymmetry. *The Accounting Review* 64 (July): 468–487.

———, J. Davis, and W. Jung. 2000. Taxpayer disclosure and penalty laws. *Journal of Public Economic Theory* 2 (2): 243–272.

Blouin, J., C. Gleason, L. Mills, and S. Sikes. 2007. What can we learn about uncertain tax benefits from FIN 48? *National Tax Journal* 60 (September): 521–35.

———, ———, ———, and ———. 2010. Pre-empting disclosure? Firms' decisions prior to FIN No. 48. *The Accounting Review* 85 (3): 791–815.

Bureau of National Affairs (BNA). 2006. *Special Report on FASB Interpretation No. 48 (FIN 48)—Accounting for Uncertainty in Income Taxes*. Accounting Policy and Practice Portfolio Series. September 21. Arlington, VA: BNA.

Dunbar, A., J. Phillips, and G. Plesko. 2009. The effect of FIN 48 on firms' tax reporting behavior. Working paper, University of Connecticut.

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes*. FASB Interpretation No. 48. Norwalk, CT: FASB.

Gleason, C., and L. Mills. 2002. Materiality and contingent tax liability reporting. *The Accounting Review* 77 (2): 317–342.

Graetz, M. J., J. F. Reinganum, and L. L. Wilde. 1986. The tax compliance game: Toward an interactive theory of law enforcement. *Journal of Law, Economics and Organization* 2 (1): 1–32.

Gupta, S., L. Mills, and E. Towery. 2009. Did FIN 48 arrest the trend in multistate tax avoidance? Working paper, University of Texas.

Lisowsky, P., L. Robinson, and A. Schmidt. 2010. Do Publicly Disclosed Tax Reserves Tell Us About Privately Disclosed Tax Shelter Activity? Working paper, University of Illinois.

Magro, A., and S. Nutter. 2009. Evaluating the strength of tax authorities: How experience affects the assessment and combination of source and relevance. Working paper, George Mason University.

Mills, L., and R. Sansing. 2000. Strategic tax and financial reporting decisions: Theory and evidence. *Contemporary Accounting Research* 17 (Spring): 85–106.

Mulligan, T. 2007. Tribune nearing IRS settlement. *Los Angeles Times* (June 2).

Pauly, E. 2008. Raising penalty standard "more likely than not" creates conflict between tax lawyers and their clients. *Washburn Law Journal* 47: 757–783.

Robinson, L. 2010. Do firms incur costs to avoid reducing pre-tax earnings? Evidence from the accounting for low-income housing tax credits? *The Accounting Review* 85 (2): 637–669.

Sansing, R. 1993. Information acquisition in a tax compliance game. *The Accounting Review* 68 (October): 874–884.

Spatt, C. 2007. Chief Economist and Director of the Office of Economic Analysis speech: The economics of FIN 48: Accounting for uncertainty in income taxes. March 8. Washington, D.C.: SEC.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1743–1761

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1743

Effects of Technical Department's Advice, Quality Assessment Standards, and Client Justifications on Auditors' Propensity to Accept Client-Preferred Accounting Methods

Terence Bu-Peow Ng
Premila Gowri Shankar
Nanyang Technological University

ABSTRACT: We conduct an experiment to investigate whether the effects of an SAS No. 90-like quality assessment standard (QAS) on auditors' decision to accept a client-preferred accounting method is jointly moderated by the nature of advice from an audit firm's technical department and the strength of the client's justification. Results indicate that, in the absence of advice, the presence of QAS does not affect auditors' decision, regardless of client justification strength. However, in the presence of advice from the technical department, the presence of QAS significantly reduces auditors' propensity to accept the client-preferred method, but only when the advice explicitly recommends the use of the most appropriate method and the client's justification is strong. These findings demonstrate the complementary roles that professional standards and audit firms' technical departments play in enhancing the quality of auditors' decisions, and indicate that the nature of advice matters.

Keywords: *advice; quality assessment; justification; auditor judgment.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

Improving the quality of accounting information provided to market participants has been a key focus of the accounting profession and regulators in recent years, more so following several widely publicized accounting scandals and audit failures. Among other measures, standards have been introduced, both in the U.S. (e.g., SAS No. 90, AICPA 1999) and internationally (e.g., ISA 260, IAASB 2009a), to require auditors to discuss with listed companies' audit committees their judgments about the appropriateness or quality, not just acceptability, of the accounting

We thank Steven Kachelmeier (senior editor), two anonymous reviewers, Dan Dhaliwal, Ted Mock, Hun-Tong Tan, Sally Wright, and participants of the 2007 International Symposium on Auditing Research (Shanghai) for helpful comments, Caroline Lee, Peter Lee, Chandrika Manohara, Cheng Heng Tan, and Pearl Tan for assistance with pilot tests and data collection, and the participants and their firms for their generous support.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: February 2007
Accepted: January 2010
Published Online: August 2010

principles and estimates in audited financial statements. Implicit in this requirement is the expectation that, while there may be different methods to account for a transaction under generally accepted accounting principles (GAAP), to enhance the quality of financial reporting, the most appropriate accounting method that best reflects the economic substance of the transaction should be used. However, prior research suggests that despite such a requirement, auditors continue to accept client-preferred accounting methods even when they may not be the most appropriate (e.g., Kadous et al. 2003; Shafer et al. 2004). Thus, understanding factors that affect auditors' decisions on clients' accounting methods remains an important research question.

The theory of motivated reasoning suggests that individuals are prone to biased reasoning in favor of a directional goal, subject to "reasonableness constraints" (Kunda 1990). In this study, we investigate whether the effect of an SAS No. 90-like standard requiring auditor assessment of the quality, not just acceptability, of a client's accounting methods (hereafter referred to as "quality assessment standard" or "QAS") on auditors' decisions is contingent on the presence and/or nature of advice and the strength of the client's justification for its preferred accounting method. We examine these two factors, as they could impose varying degrees of "reasonableness constraints" to counter auditors' directional goal of supporting client-preferred accounting methods. Providing evidence on this issue sheds light on the boundary conditions under which prior findings on the efficacy of quality assessment standards hold (e.g., Kadous et al. 2003), as well as how these factors interact to jointly influence auditors' proclivity to support client-preferred accounting methods.

Consultation with other experts within the audit firm, including personnel from the firm's technical department, constitutes an important part of an audit firm's quality control mechanisms (e.g., AICPA 2006, 2009; IAASB 2009b) and is widely used for resolving ambiguous or contentious accounting issues encountered by auditors (e.g., Gibbins and Emby 1985). However, prior studies have not systematically investigated how the presence or nature of advice received by auditors affects their decisions, either independently or jointly with other features in the audit environment.[1] Research in psychology suggests that, although incorporating advice from others generally increases decision accuracy, individuals often underutilize or discount the advice received (see Bonaccio and Dalal [2006] for a review of this literature). Thus, providing evidence on the extent to which the presence or nature of advice moderates the effect of QAS, by either tightening or loosening the reasonableness constraints, contributes to a better understanding of the role and limitations of consultations in enhancing auditors' decisions.

In addition, we investigate whether the strength of the client's justification for its preferred accounting method further moderates the joint effects of QAS and advice, given that firms with incentives to adopt aggressive accounting methods would likely attempt (or be required by their auditors) to justify their accounting method choices. To the extent that the reasonableness constraints on auditors' proclivity to support client-preferred methods are loosened by strong justifications provided by the client, the incremental role and impact of QAS and advice may increase.

To provide evidence on the above issues, we conduct a two-stage experiment with auditors from two Big 4 firms in Singapore, where SAS No. 90-like standards have not been implemented at the time of the experiment.[2] In the first stage, participants decide whether to accept a client-

[1] Salterio (1996) and Salterio and Koonce (1997) examine a context where auditors need to make their own judgment about the appropriateness of client's accounting treatments based on precedents received from the firm's central research unit (i.e., no specific advice on the appropriateness of the accounting issue is provided). Perkins (2003) examines the effects of accountability type and justifiee preference on tax accountants' informal consultation and documentation stylization behaviors, but not how advice-recipients' decisions are influenced by advice received.

[2] We employ a two-stage experiment to capture participants' pre- and post-advice decisions so as to assess how auditors' decisions are influenced by the advice provided. Bonaccio and Dalal (2006) argue that such a design should be used,

preferred accounting method relating to a new sales contract in the absence of any advice. We manipulate, between-subjects, the presence of QAS and the strength of the client's justification for its preferred accounting method. This stage allows us to determine the effects of QAS and client justification strength in the absence of advice, and provides a baseline for assessing the incremental effects of the advice subsequently provided.

In the second stage, participants make the same decision on whether to accept the client-preferred method after receiving advice from the audit firm's technical department.[3] The advice reduces the ambiguity surrounding the accounting issue by clarifying that, while the client-preferred method is an acceptable method under GAAP, an alternative, less aggressive method is the most appropriate method. This stage allows us to examine the extent to which auditors would: (1) continue to support the client-preferred method, given that it has been identified by the technical department to be an acceptable, albeit not the most appropriate, method under GAAP; or (2) reject the client-preferred method in favor of the most appropriate method identified by the technical department.[4,5] In addition, we manipulate the nature of the advice, by either including or not including an explicit recommendation to use the most appropriate method, to investigate whether providing an explicit recommendation has an incremental effect in reducing auditors' propensity to support the client-preferred method (as it further tightens the reasonableness constraints) beyond information provided in the advice regarding the relative quality of the client's method.[6]

Results indicate that, in the absence of any advice from the technical department, auditors' propensity to accept the client-preferred method is significantly influenced by the strength of the client's justification but not by the presence of QAS. However, in the presence of the technical department's advice indicating that an alternative method is the most appropriate method, we find a significant three-way interaction between presence of QAS, nature of the advice, and client justification strength. Specifically, when the client's justification is weak, auditors' propensity to accept the client-preferred method is not affected by the presence of QAS, regardless of the nature of the technical department's advice. However, when the client's justification is strong, auditors' propensity to accept the client-preferred method decreases in the presence of QAS, but only when

---

where possible, to facilitate the evaluation of advice utilization. Importantly, this design replicates what often happens in practice. Discussions with the participating firms' technical department staff indicate that auditors who are seeking consultation on an issue are required, wherever possible, to state their preliminary position on the issue being consulted. Hence, auditors generally make their judgments twice, both before and after receiving advice from the technical department.

[3] Although auditors could potentially seek and receive advice from various sources, we examine the context where advice is sought and provided by the firm's technical department given that prior research indicates that audit firms' technical departments serve as an important source of information and consultation for auditors (Gibbins and Emby 1985; Danos et al. 1989; Salterio and Denham 1997). The audit firm's technical department may also be referred to as the accounting consultation unit, professional practice department, professional standards group, etc. Salterio and Denham (1997) describe the common structure and features of these technical departments.

[4] We do not examine the context where the advice indicates that the client-preferred method is not an acceptable method under GAAP, since we expect few, if any, auditors would accept an accounting method that clearly violates GAAP. Moreover, this context is less germane for testing the efficacy of QAS, since the intent of such a standard is to encourage auditors to consider the appropriateness, *not just acceptability*, of clients' accounting methods, particularly in situations where there are alternative methods that are acceptable under GAAP.

[5] Our study examines a context where advice issued by the audit firm's technical department is non-binding (e.g., Kennedy et al. 1997). Based on our discussions with the technical department staff of the participating firms, in general, the technical department personnel provide their opinion on the accounting issue under consultation, and leave the final decision to the discretion of the audit engagement partner.

[6] The technical department staff of both the participating firms indicate that in situations where there are more than one accounting methods that are acceptable under GAAP for a transaction undertaken by the client, the technical department would, when consulted on the issue, communicate the acceptability of the client-preferred method under GAAP, identify the method deemed most appropriate, and frequently *but do not always* explicitly recommend/suggest that the client should adopt the most appropriate accounting method.

the advice includes an explicit recommendation that the client should adopt the most appropriate method. These results suggest that the effect of QAS is jointly moderated by the nature of advice received and the strength of the client's justification.

Our study contributes to the literature in several ways. First, our study adds to an important stream of research that investigates factors affecting auditors' propensity to support client-preferred accounting methods (e.g., Trompeter 1994; Hackenbrack and Nelson 1996; Salterio 1996; Salterio and Koonce 1997; Jenkins and Haynes 2003; Kadous et al. 2003) by demonstrating how such decisions can be jointly influenced by both intra- and extra-firm quality control mechanisms as well as the strength of clients' justifications. More specifically, our study extends prior research on the efficacy of SAS No. 90-like quality assessment standards (Kadous et al. 2003) by providing evidence on factors that could moderate their effects. Our finding that such a standard reduces auditors' propensity to accept a client-preferred, strongly justified method only when the advice from the technical department contains an explicit recommendation demonstrates the complementary roles that professional standards and audit firms' technical departments play in enhancing auditors' decisions. Second, our results show that merely reducing the ambiguity of the accounting issue by clarifying the relative quality of client-preferred methods (*vis-à-vis* other alternative methods) may not be sufficient to mitigate auditors' propensity to accept less appropriate client-preferred methods, even in the presence of QAS. Additional steps such as the technical department explicitly recommending (or advocating) the use of the most appropriate methods to enhance financial reporting quality may be necessary. This finding has important implications for audit firms in terms of the nature of advice communicated by their technical departments, as well as for regulators in terms of the potential need to further clarify auditors' and audit committees' responsibility for promoting the use of the most appropriate accounting methods. Last, we add to the broader literature on the impacts of advice on individuals' judgments and decisions (e.g., Sniezek and Buckley 1995; Yaniv 2004; Bonaccio and Dalal 2006) by demonstrating how the influence of advice on professionals' decisions is contingent on its nature, strength of justifications provided by others, as well as the presence of external constraints (e.g., professional standards).

Section II presents our hypothesis development. Section III follows with a description of the research methods used, and Section IV presents the results. Section V concludes with a discussion of the study's main findings, limitations, and opportunities for future research.

II. HYPOTHESIS DEVELOPMENT

A number of prior studies have documented auditors' proclivity to support client-preferred accounting methods within reasonable limits (e.g., Trompeter 1994; Hackenbrack and Nelson 1996; Salterio 1996; Salterio and Koonce 1997; Jenkins and Haynes 2003; Kadous et al. 2003). In the study most closely related to ours, Kadous et al. (2003) find that the presence of a quality assessment standard (QAS) does not mitigate auditors' propensity to support an aggressive method preferred by the client. Instead, it exacerbates the influence of directional goal commitment, an individual pro-client motivational factor, on auditors' decisions. The theory of motivated reasoning suggests that auditors with pro-client directional goals are likely to support their client-preferred method within reasonableness constraints—that is, as far as the decisions are justifiable (Kunda 1990). One would expect the presence of QAS (which requires auditors' assessment and audit committee communication of the quality of clients' accounting methods) to impose a reasonableness constraint, thereby making it less justifiable for auditors to support the client-preferred method. However, this effect is not observed in Kadous et al. (2003), which we surmise could be due to one of the following reasons: (1) auditors may not be aware, in the absence of any advice, that there is a more appropriate method, particularly given the ambiguity of the accounting issue; and (2) even if auditors are aware that there is a more appropriate method, they may continue to

support the client-preferred method because a QAS-like standard does not explicitly mandate auditors to ensure that the most appropriate accounting method is adopted by the client. Indeed, recent survey research in the U.S. suggests that, even in the post-SAS No. 90 environment, auditors indicate that they may not insist on the use of the most appropriate method by the client, as long as the client-preferred method is acceptable under GAAP (Shafer et al. 2004). Thus, even in the presence of QAS, auditors may continue to support the client-preferred method.

Kadous et al. (2003) examine a context in which no advice on the relative quality of the client's accounting method is available, and the client provides a relatively weak justification for their preferred method based on simplicity of the method and a desire to manage earnings expectations (see Kadous et al. 2003, 768). In practice, auditors are likely to seek advice from other experts to help resolve ambiguous and/or contentious accounting issues (e.g., Gibbins and Emby 1985). Moreover, a client with incentives to adopt aggressive accounting methods is likely to provide strong (rather than weak) justifications in support of its preferred method. These factors could independently or jointly moderate the influence of QAS on auditors' propensity to support client-preferred methods by either tightening or loosening the reasonableness constraints on auditors' proclivity to support client-preferred methods. Our study sheds light on these issues by examining whether the effect of QAS is greater in the presence of advice from the audit firm's technical department on the relative quality of the client's preferred method, and whether this moderating effect of advice is contingent on the nature of the advice and the strength of the client's justification for its preferred method.

Suppose that auditors seek and receive advice from their firm's technical department indicating that although the client-preferred method is an acceptable method under GAAP, an alternative (less aggressive) accounting method is the most appropriate method.[7] The indication of the most appropriate method by the technical department reduces the ambiguity of the accounting issue by providing auditors with a benchmark to evaluate against the more aggressive client-preferred method. This should tighten the reasonableness constraints on auditors' proclivity to support the client-preferred method by reducing the justifiability of accepting a less appropriate method preferred by the client. To what extent will the presence of the advice enhance the effect of QAS in reducing auditors' propensity to support the client-preferred method?

Prior research suggests that individuals use advice to enhance the justifiability of their decision or share responsibility with others for their decision rather than enhancing decision quality *per se* (Harvey and Fischer 1997; Kennedy et al. 1997), and that even advice from experts may not be optimally used by advice-recipients due to egocentric discounting (e.g., Yaniv and Kleinberger 2000; Bonaccio and Dalal 2006). Hence, even in the presence of QAS, auditors could selectively or strategically use information in the advice (that the client-preferred method is an acceptable, albeit not the most appropriate, method under GAAP) to support the client-preferred method, given a pro-client directional goal. On the other hand, if the advice from the technical department incorporates an explicit recommendation that the client should use the most appropriate method, this should further tighten the reasonableness constraints on auditors' motivated reasoning by making it even less justifiable to accept a less appropriate client-preferred method, since doing so would imply ignoring the explicit recommendation provided by the firm's technical department. Thus, the presence of the technical department's advice is more likely to enhance the effect of QAS in reducing auditors' propensity to support the client-preferred method when that advice contains an explicit recommendation to use the most appropriate method than when it does not.

[7] Consistent with the general advice literature (see Gibbons 2003; Bonaccio and Dalal 2006), we use the term "advice" broadly to include one that provides information about different alternatives, suggests new alternatives, and/or recommends a specific course of action (Gibbons 2003; Bonaccio and Dalal 2006). Thus, advice may or may not contain an explicit recommendation.

However, we argue that the moderating effect of advice nature on the effect of QAS discussed above is likely to be greater (or more observable) when the client's justification for its preferred method is strong than when it is weak. Psychology research suggests that when people are motivated to evaluate an externally provided message, they will carefully scrutinize the arguments in the message to arrive at their own conclusions (Petty and Cacioppo 1986). This reasoning suggests that the strength of the client's justification for its preferred accounting method is likely to positively influence auditors' assessments of the acceptability and appropriateness of the method, thereby increasing their perceived justifiability of supporting the method. Further, psychological research in motivated reasoning (e.g., Kunda 1990) and elastic justification (e.g., Hsee 1995) suggests that, given pro-client incentives and ambiguity of the accounting issue (e.g., Hackenbrack and Nelson 1996; Salterio and Koonce 1997; Kadous et al. 2003), auditors are likely to use strong justifications provided by the client to bolster their support for the client-preferred method. Thus, a strong (weak) client justification likely loosens (tightens) the reasonableness constraints on auditors' proclivity to accept client-preferred methods (Kunda 1990; Hsee 1995), by enhancing (reducing) the justifiability of supporting the client-preferred method. Consequently, any incremental effect of QAS or advice in reducing the propensity of auditors to support client-preferred methods, through the tightening of reasonableness constraints, is likely greater (or more observable) in the presence of a strong client justification than a weak client justification.[8] Taken together, the above discussions suggest that the effect of QAS in reducing auditors' propensity to accept the client-preferred method will be greater when the advice from the technical department contains an explicit recommendation *and* when the client justification is strong. Our hypothesis on how the effect of QAS is jointly moderated by advice nature and client justification strength is formally stated as follows (Figure 1, Panel A, illustrates our expectations):

Hypothesis: Given the technical department's advice that an alternative accounting method (other than the client's preferred method) is the most appropriate method, the presence of a quality assessment standard will reduce auditors' propensity to accept the client's preferred method, particularly when the advice explicitly recommends the use of the most appropriate method and the client has provided a strong justification for its preferred method.

III. METHOD

Design

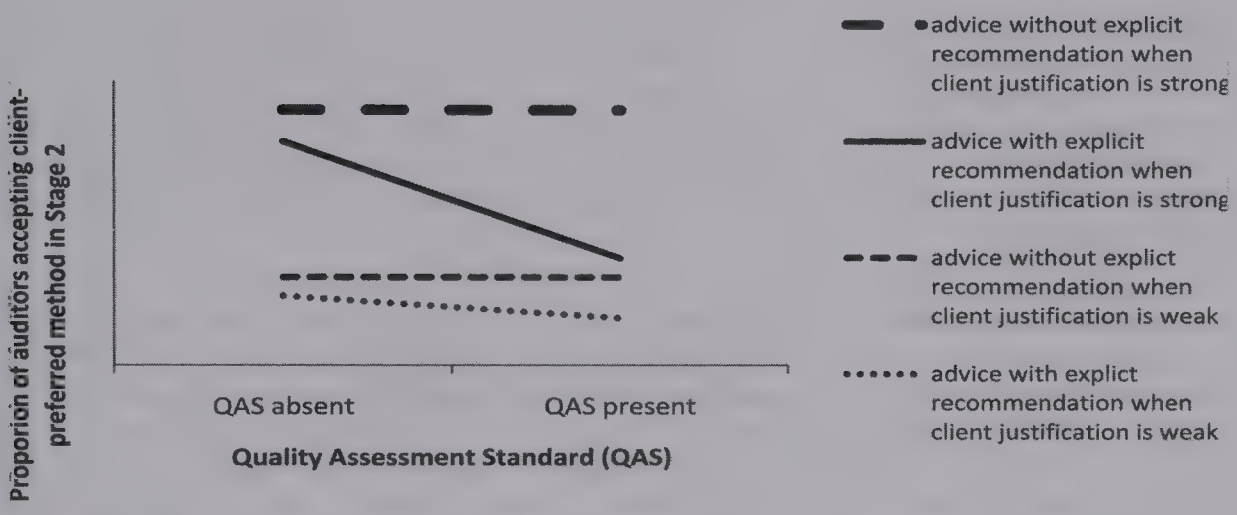
We conduct the experiment in two stages. In stage 1 (in which no advice is provided), we employ a 2 × 2 between-subjects design, with QAS (absent or present) and client justification strength (weak or strong) as manipulated factors. In stage 2 (in which participants are provided with hypothetical advice from the firm's technical department), we employ a 2 × 2 × 2 between-subjects design, with QAS, client justification strength, and advice nature (with or without an explicit recommendation) as manipulated factors. Our experimental case is adapted from that of Kadous et al. (2003).[9] The case involves an ambiguous revenue recognition issue pertaining to a new sales contract. The client prefers to use the straight-line method to recognize the revenue evenly over the contract period, which will enable it to meet its earnings projection for the year

[8] Kadous et al. (2003) show that even in the presence of a weak client justification, auditors are still susceptible to the influence of their directional goal commitment. Thus, although we argue that the expected joint effects of QAS and advice nature is likely smaller when the client's justification is weak (versus strong), our argument does not preclude the possibility that auditors may still be influenced by their directional goals to support client-preferred methods in the presence of a weak client justification, as further demonstrated by the results of our further analyses, reported later.

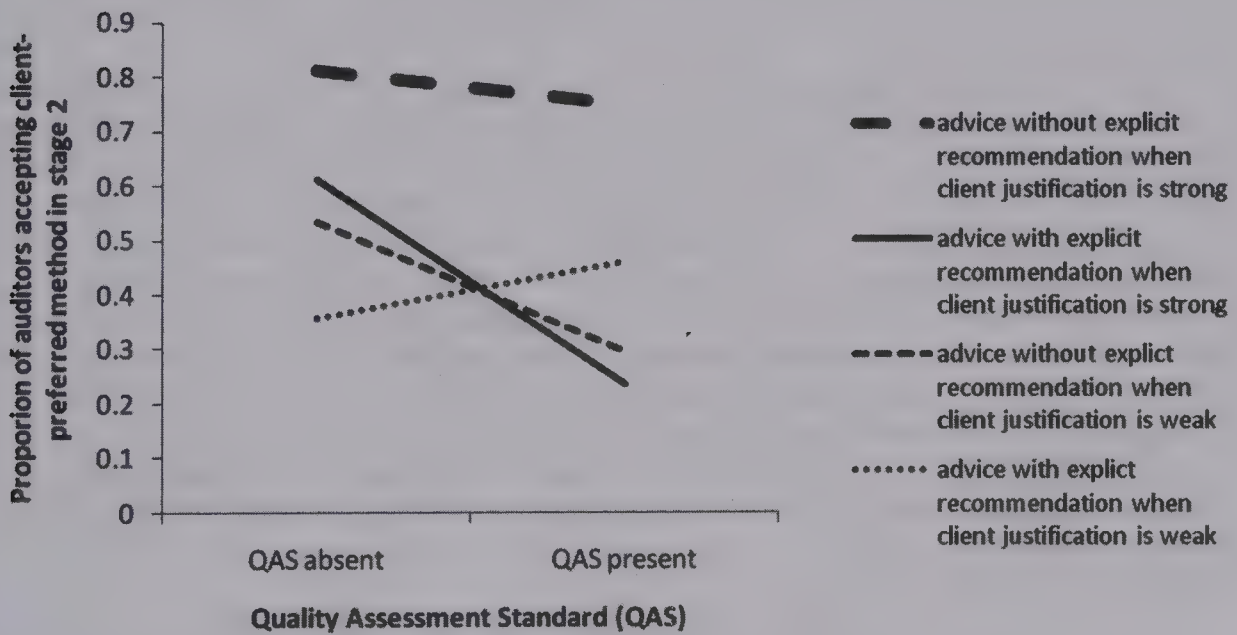
[9] We thank Kathryn Kadous for making the case available to us.

FIGURE 1
Prediction and Findings on the Effects of QAS, Advice Nature, and Client Justification Strength on Auditors' Propensity to Accept Client-Preferred Method (in Stage 2)

Panel A: Predicted Effects of QAS, Advice Nature, and Client Justification Strength on Auditor's Stage 2 Decisions



Panel B: Observed Effects of QAS, Advice Nature, and Client Justification Strength on Auditor's Stage 2 Decisions



and maintain an increasing earnings trend. However, this method is potentially aggressive relative to an alternative method based on actual usage. An average of 32 percent of the participating auditors in Kadous et al.'s (2003) study accepted the client-preferred straight-line method across various experimental conditions, suggesting that there is ambiguity (and doubt) surrounding the appropriateness of this method.

We manipulate the presence of QAS by providing half the participants with instructions similar to the requirements underlying SAS No. 90. Specifically, these participants were asked to assume that the auditing standards require them to discuss with the audit committee the quality (that is, the appropriateness), not just the acceptability, of the client's accounting principles and underlying estimates. Further, they were told that "the quality (that is, the appropriateness) of an accounting method is a function of the extent to which it is consistent with the conceptual framework and accounting standards, and the degree to which it captures the economic substance of the transaction under consideration." In addition, they were asked to rate the appropriateness of the client's method of recognizing revenue for the contract (i.e., the straight-line method) prior to deciding whether to accept the client's method. These instructions were repeated in the second stage of the experiment, in which participants were provided with advice from the technical department (described below). For the other half of the participants, no such instructions were provided in either stage of the experiment. Instead, these participants were asked to decide whether to accept the client's use of the straight-line method without first assessing the appropriateness of the method.

We manipulate the nature of the advice provided by the technical department in terms of whether it contains an explicit recommendation that the most appropriate method should be used. All participants were told to assume that their firm's technical department has been consulted on the issue. They were provided a summary of the key points communicated by the technical department, which describes four alternative methods for recognizing the contract revenue and indicates that, while the client's preferred method is an acceptable method under GAAP, the actual usage rate method is the most appropriate method. For participants in the "advice with explicit recommendation" condition, the technical department's communication included a recommendation that, although the straight-line method is acceptable under GAAP, to enhance financial reporting quality, the client should adopt the most appropriate method (i.e., the actual usage rate method). This recommendation was not provided to participants in the "advice without explicit recommendation" condition.

Last, we manipulate the strength of the client's justification for the use of the straight-line method by varying the content of the justification provided by the client's Chief Financial Officer (CFO). Specifically, for the weak client justification condition, the CFO argues that the straight-line method is appropriate for recognizing revenue from the contract, as it is easy to apply and it generates a smooth income stream. This justification is similar to that used in Kadous et al. (2003). In contrast, for the strong client justification condition, the CFO argues that the method is appropriate based on past consumption patterns that suggest that the overall usage over the contract period may not exceed the minimum usage that the customer has agreed to pay. The latter justification is stronger than the former, since it is based on an analysis of the transaction rather than on simplicity or income-smoothing objectives.[10]

Our main dependent variable of interest is auditors' final (stage 2) decision on whether to accept the client's use of the straight-line method for recognizing the contract revenue, given the technical department's advice.[11]

Administration and Procedures

The experiment was administered in a controlled setting during the participating firms' training sessions. In stage 1, participants were randomly assigned to each of the four client QAS/client

[10] Research by Bame-Aldred and Kida (2007, 506) indicates that auditors do not consider justifications such as avoiding variability in the earnings stream to be as important as the client does for supporting the use of a revenue recognition method. We developed our strong justification based on the response of one of the pilot-test participants who teaches advanced financial accounting and is a former senior audit manager.

[11] Although our focus is on auditors' final decision in the presence of the technical department's advice, we also analyze the change in auditors' decisions between stages. We report the related results as part of further analyses.

justification strength treatments. In stage 2, we hold constant the client justification strength/QAS treatment, and participants within each justification strength/QAS treatment were randomly assigned to one of the two nature of advice treatments. The experimental materials for the two experimental stages were enclosed in separate envelopes. At the start of the experiment, the participants were provided with a brief introduction to the project, including an assurance of confidentiality, and were asked to provide their name to acknowledge participation. They then proceeded to read the experimental case relating to a hypothetical client that contains our manipulation of QAS and client justification strength, as described above. Following this instruction, participants were asked to make several judgments on the case, including whether to accept the client-preferred method.

Upon completing stage 1, participants were asked to place the completed case materials back into the envelope provided and proceed to the next envelope. They were first asked to complete a short questionnaire eliciting their assessments of the complexity of the accounting issue, engagement risk of the client, and their knowledge on the accounting issue, as well as some demographic information. They then proceeded to perform the stage 2 experimental task by reading a hypothetical advice memorandum received from their firm's technical department and making similar judgments on the case as in stage 1. After completing the stage 2 experimental task, they completed a debriefing questionnaire (enclosed in a separate envelope), which includes manipulation check and other questions.[12]

Participants

Participants were 129 experienced audit seniors, supervisors, and managers from two Big 4 firms in Singapore. Three participants were eliminated due to incomplete responses, yielding a final sample of 126 auditors, comprising 66 audit seniors (with an average audit experience of 3.6 years, ranging from 3 to 8 years) and 59 audit supervisors and managers (with an average audit experience of 5.5 years, ranging from 4 to 16 years).[13] The majority (80 percent) of the participants indicated experience auditing listed companies, and about half of them (51 percent) participated in, or prepared for, discussions with clients' audit committees. On average, over the past two years, participants have consulted the technical department personnel on clients' accounting issues on a formal basis (i.e., by submitting a formal written request) in 29 percent of their audit engagements, and on an informal basis (such as over the phone) in 49 percent of their audit engagements. They have also consulted other audit colleagues in the firm (outside the audit engagement team and the technical department) on clients' accounting issues in 64 percent of their audit engagements, on average, over the past two years. None of the above experience measures vary significantly between the experimental conditions at the 5 percent level. Our results do not change when we exclude participants who do not have experience in listed company audits, nor when we include any of the other experience measures as a covariate in our analyses. None of these experience measures are significantly associated with the dependent variable at the 5 percent level.

[12] We handed out the three envelopes all at once, with the first envelope placed on top. At the start of the experiment, we verbally instructed participants to open the three envelopes (labeled A, B, and C) in sequence (i.e., to proceed to envelope B only after they have completed the materials in envelope A and placed the completed responses back into envelope A, and so on). The researchers also moved around the training room to observe whether participants followed the instruction, and no violations were observed.

[13] One participant did not indicate his or her rank. Similar results for hypothesis tests are obtained when this participant is eliminated from the analyses. The hypothesis test results do not vary significantly by firm or rank at the 5 percent level. Unless otherwise stated, two-tailed p-values are reported. One-tailed p-values are reported for directional tests.

IV. RESULTS

Manipulation and Other Checks

We did not conduct a manipulation check on the quality assessment standard (QAS) manipulation, given that participants were either told or not told about the standard's requirements (i.e., to assess and communicate with the audit committee the quality of the client's accounting methods), and only participants in the condition where QAS is present were asked to assess the appropriateness of the client's preferred method. To the extent that participants assess the quality of the client's accounting method even without being asked to do so, our tests should be biased against finding an effect of QAS (see also Kadous et al. 2003, 769, footnote 8). As a check on our manipulation of the nature of the firm's advice in stage 2, we ask participants to rate the extent to which the technical department explicitly recommends that the client should adopt the most appropriate method to enhance financial reporting quality (1 = very low extent, 7 = very high extent). Participants' ratings are higher in the "advice with explicit recommendation" condition than in the "advice without explicit recommendation" condition (means = 5.1 and 4.6, respectively; $p = 0.038$), suggesting that they attend to our manipulation of advice nature.[14] Participants rate the client's justification to be stronger (means = 4.4 and 3.3, respectively, $p < 0.001$; 1 = very weak; 7 = very strong) and more conceptually sound (means = 4.3 and 3.4 respectively; $p < 0.001$; 1 = very low extent; 7 = very high extent) in the strong client justification condition than in the weak client justification condition. Thus, our manipulation of client justification strength appears successful.[15]

We also ask participants to indicate the extent to which they agree with the following statement: "When there is more than one method of accounting for a transaction that could be deemed acceptable under GAAP, auditors should require the client to use the most appropriate accounting method" (1 = disagree completely, 7 = agree completely). Participants' mean agreement rating is significantly higher than neutral (mean = 5.4; $p < 0.001$), and does not vary by nature of advice or other experimental conditions (all $p \geq 0.245$). Thus, whether the advice contains an explicit recommendation does not appear to influence auditors' belief that they should ensure the use of the most appropriate method by the client, although we predict and find (as reported later) that auditors' decisions are influenced by advice nature. This disconnect between what professionals believe they should do and what they actually do has been demonstrated in prior research (e.g., Libby and Tan 1999).

Recall that participants in the condition in which QAS is present are asked to assess the appropriateness of the client-preferred method prior to deciding whether to accept it. Consistent with our expectation that client justification strength will positively influence auditors' assessment of the appropriateness of the client-preferred method, participants' appropriateness ratings for the client-preferred method are significantly higher when client justification is strong than when it is weak, in both stage 1 (means = 0.46 and 0.33, respectively, with 0 percent = not at all appro-

[14] We ask a subset of participants (n = 105) to indicate the most appropriate method identified by the technical department and whether the technical department assessed the client-preferred straight-line method to be an acceptable method under GAAP. The majority (n = 86 or 82 percent) answered both questions correctly, and we obtain similar results for the tests of hypotheses based on these 86 participants. These participants generally agree with the technical department's choice of the most appropriate method (mean = 5.6;1 = disagree completely; 7 = agree completely), although their agreement ratings are higher when client justification is weaker than when it is stronger (means = 5.9 and 5.3, respectively; $p = 0.019$). Neither the nature of the advice nor the presence of QAS has any significant impact on the agreement ratings at the 5 percent level.

[15] Participants do not have differential beliefs about the extent to which the client's justification is based on the desire to manage earnings across the client justification strength conditions (means = 4.7 for both conditions; $p > 0.500$; 1 = very low extent; 7 = very high extent). Thus, our manipulation of client justification strength does not result in differential assessments of the client's incentive to manage earnings, which could otherwise affect auditors' reliance on management representations (see Anderson et al. 2004).

priate, 100 percent = completely appropriate; p = 0.025) and stage 2 (means = 0.53 and 0.40, respectively; p = 0.015).[16] Advice nature does not significantly affect auditors' stage 2 appropriateness assessments, whether independently or jointly with client justification strength (all p ≥ 0.154). We also ask a subset of participants in all experimental conditions in stage 1 whether they believe that the client-preferred method is acceptable under GAAP.[17] A higher proportion of participants answer "yes" when the client's justification is strong than when it is weak (proportions = 0.50 and 0.27, respectively; p = 0.012). These proportions do not vary significantly by QAS (all p ≥ 0.342). Overall, the results indicate that client justification strength influences auditors' assessments of both the acceptability and appropriateness of the client-preferred method as expected.

Tests of Hypothesis

Our research hypothesis predicts that, given the technical department's advice that an alternative accounting method (other than the client's preferred method) is the most appropriate method, the presence of a quality assessment standard will reduce auditors' propensity to accept the client's preferred method, particularly when the advice explicitly recommends the use of the most appropriate method and the client has provided a strong justification for its preferred method. We first analyze the effects of QAS and client justification strength on auditors' decisions in the absence of advice. Table 1, Panel A presents the proportion of auditors who accept the client-preferred straight-line method (hereafter, P_{Accept}) by experimental treatment in stage 1, in which advice is absent. Panel B of Table 1 presents the categorical ANOVA results, with QAS and client justification strength as independent variables, and the auditors' stage 1 decision as the dependent variable. Results indicate a significant main effect of client justification strength, with a higher P_{Accept} when the client's justification is strong than when it is weak (P_{Accept} = 0.43 and 0.15, respectively; p < 0.001). Neither the main effect of QAS nor the interaction effect between client justification strength and QAS is statistically significant (all p ≥ 0.389).[18] These results indicate that in the absence of any advice, auditors' propensity to accept the client-preferred method is significantly influenced by the strength of the client's justification, but not by the presence of QAS. The latter result is consistent with Kadous et al. (2003).[19]

Table 2, Panel A presents the P_{Accept} by experimental treatment in stage 2, in which advice is provided. Panel B of Table 2 presents the categorical ANOVA results, with QAS, client justification strength, and advice nature as independent variables, and the auditors' stage 2 decision as the dependent variable. Results indicate a statistically significant three-way interaction (p = 0.047).[20] Tests of the simple effect of QAS within each of the advice nature/client justification strength conditions indicate that P_{Accept} is significantly lower in the presence of QAS than in its absence

[16] Eight participants in the "QAS present" condition did not provide their appropriateness assessments in stage 1, presumably because of the greater ambiguity of the accounting issue in the absence of any advice. Inferences relating to our tests of the hypothesis remain the same when these participants are excluded.

[17] This question was inserted only after the first training session. Hence, the analysis is based on 102 participants who responded to the question. We did not ask this question in stage 2 of the experiment given that the advice provided already indicates that the client-preferred method is an acceptable method under GAAP.

[18] Including advice nature as a third independent variable in the analysis does not change these results. Advice nature does not significantly affect, nor interact with the other independent variables to jointly affect, auditors' decisions (all p ≥ 0.196).

[19] We also elicited, at the end of stage 1, participants' assessments of their likelihood of consulting (1) their firm's technical department personnel (mean = 0.67), and (2) other colleagues outside the audit team and technical department (mean = 0.69), on the accounting issue. These assessments do not vary significantly by QAS or client justification strength (all p ≥ 0.188).

[20] The three-way interaction remains marginally significant (p = 0.069) when we replace our manipulated client justification strength variable with the participants' perceived strength of client's justification using a dichotomous variable based on a median split.

TABLE 1

Effects of Quality Assessment Standard and Client Justification Strength on Auditors' Decision in the Absence of Technical Department's Advice (Stage 1)

Panel A: Descriptive Statistics on the Proportion of Auditors Accepting the Client-Preferred Method by Experimental Treatment

Client Justification Strength[b]	Quality Assessment Standard[a]		
	Absent	Present	Average
Weak	0.14 (n = 29)	0.17 (n = 30)	0.15 (n = 59)
Strong	0.38 (n = 34)	0.48 (n = 33)	0.43 (n = 67)
Average	0.27 (n = 63)	0.33 (n = 63)	0.30 (n = 126)

Panel B: Categorical ANOVA Results

Effects	df	χ^2	p
Client Justification Strength (CJS)	1	13.62	<0.001
Quality Assessment Standard (QAS)	1	0.74	0.389
CJS × QAS	1	0.23	0.629

[a] Quality assessment standard (QAS) is manipulated by providing half the participants with instructions similar to the requirements underlying SAS No. 90. Participants in the QAS present condition are asked to assume that the auditing standards require them to discuss with the audit committee the quality (that is, the appropriateness), not just the acceptability, of the client's accounting principles and underlying estimates, and to rate the appropriateness of the client's method of recognizing revenue for the contract prior to deciding whether to accept the client's method. These instructions are not provided to participants in the QAS absent condition.

[b] Client justification strength (CJS) is manipulated by varying the content of the justification provided by the client's CFO. For the weak client justification condition, the CFO argues that the straight-line method is appropriate for recognizing revenue from the contract as it is easy to apply and generates a smooth income stream. For the strong client justification condition, the CFO argues that the method is appropriate based on past consumption patterns that suggest that the overall usage over the contract period may not exceed the minimum usage that the customer has agreed to pay.

TABLE 2

Effects of Advice Nature, Quality Assessment Standard, and Client Justification Strength on Auditors' Decisions in the Presence of Technical Department's Advice (Stage 2)

Panel A: Descriptive Statistics on the Proportion of Auditors Accepting the Client-Preferred Method by Experimental Treatment

Client Justification Strength	Advice Nature (Explicit Recommendation Provided?)[a]	Quality Assessment Standard		
		Absent	Present	Average
Weak	No	0.53 (n = 15)	0.29 (n = 17)	0.41 (n = 32)
	Yes	0.36 (n = 14)	0.46 (n = 13)	0.41 (n = 27)
	Average	0.45 (n = 29)	0.37 (n = 30)	0.41 (n = 59)
Strong	No	0.81 (n = 16)	0.75 (n = 16)	0.78 (n = 32)
	Yes	0.61 (n = 18)	0.24 (n = 17)	0.43 (n = 35)
	Average	0.71 (n = 34)	0.48 (n = 33)	0.60 (n = 67)

Panel B: Categorical ANOVA Results

Effects	df	χ^2	p
Quality Assessment Standard (QAS)	1	3.00	0.083
Advice Nature (ADV)	1	4.80	0.028
Client Justification Strength (CJS)	1	5.32	0.021
CJS × QAS	1	0.84	0.359
CJS × ADV	1	4.58	0.032
QAS × ADV	1	0.01	0.927
CJS × QAS × ADV	1	3.95	0.047

[a] Advice nature is manipulated by either including or not including an explicit recommendation in the technical department's advice that "although the straight-line method is acceptable under GAAP, to enhance financial reporting quality, the client should adopt the most appropriate method (i.e., the actual usage rate method)."

only when the advice contains an explicit recommendation and client justification is strong ($P_{Accept} = 0.24$ and 0.61, respectively; $p = 0.015$), but not in the other three conditions (all $p \geq 0.159$).[21] These results provide support for our research hypothesis, and suggest that both the nature of the advice and the strength of the client's justification jointly moderate the effect of QAS. Figure 1, Panel B, illustrates the results. Alternative analyses of the three-way interaction indicate that P_{Accept} is significantly lower when the advice contains an explicit recommendation than when it does not, only in the condition where QAS is present and client justification is strong ($P_{Accept} = 0.24$ and 0.75, respectively; $p = 0.001$), but not in the other three QAS/client justification strength conditions (all $p \geq 0.182$).[22]

Additional Analyses on Change in Auditors' Decisions between Stages

Our tests of hypothesis above focus on auditors' final decision on the client's accounting method given the advice from the technical department, and demonstrate how the effect of QAS is moderated by the nature of the advice and the strength of the client's justification. Here, we analyze how the presence of advice changes auditors' decisions between stages. Table 3 presents the P_{Accept} within each of the between-subject treatment conditions at both stages. As shown in Table 3, P_{Accept} increases from stage 1 to stage 2 in all conditions except when client justification is strong, QAS is present, and the advice contains an explicit recommendation. These results suggest that, with the exception of the above-mentioned condition, auditors strategically use information provided in the technical department's advice (that the client-preferred is an acceptable, albeit not the most appropriate, method under GAAP) to increase their support for the client-preferred method.

We perform a mixed categorical ANOVA that includes all three between-subject variables (i.e., QAS, advice nature, and client justification strength) and participants' stage 1 and stage 2 decisions as repeated measures (hereafter, Stage). Results (untabulated) indicate no statistically significant effects at the 5 percent level except for the main effects of Stage ($p < 0.001$), advice nature ($p = 0.041$), and client justification strength ($p = 0.001$), as well as the interaction effect between advice nature and client justification strength ($p = 0.050$) and the interaction effect between Stage and QAS ($p = 0.009$). The last result suggests that the increase in P_{Accept} from stage 1 to stage 2 is moderated by QAS.

Further analyses indicate that when client justification is weak, only the main effects of stage and client justification strength are statistically significant at the 5 percent level, suggesting that neither QAS nor advice nature moderates the increase in P_{Accept} from stage 1 to stage 2. However, when client justification is strong, there is a statistically significant interaction effect between Stage and QAS ($p = 0.005$).[23] Further tests indicate that the effect of Stage is statistically signifi-

[21] Inferences relating to our tests of hypothesis remain the same when we include auditors' pre-advice (stage 1) decision as a control variable in the analyses. We also re-perform the analyses using auditors' perceptions of the accounting method that would ultimately be used in the financial statements after discussions with the management and audit committee. Results (based on 119 participants who responded) similarly indicate that presence of QAS does not significantly reduce auditors' perception that the client's preferred method will ultimately be used except in the condition where the advice contains an explicit recommendation and client justification is strong ($p = 0.029$).

[22] We also test the joint impacts of QAS and advice nature on the influence of client justification strength on auditors' decisions. Results show that in the absence of QAS, client justification strength significantly increases P_{Accept}, regardless of advice nature ($p = 0.024$). However, in the presence of QAS, client justification strength significantly increases P_{Accept} only when advice does not contain an explicit recommendation ($p = 0.003$), but not when it contains an explicit recommendation ($p = 0.189$).

[23] Although these results suggest a significant interaction effect between client justification strength, QAS and Stage, neither the three-way interaction effect between Stage, QAS, and client justification ($p = 0.155$) nor the four-way interaction effect between Stage, QAS, advice nature, and client justification strength ($p = 0.105$) is statistically significant in the mixed categorical ANOVA reported earlier.

TABLE 3

Proportions of Auditors Accepting the Client-Preferred Method (P_{Accept}) in stage 1 and stage 2, and the Difference in P_{Accept} between Stages, for each Experimental Condition

Client Justification Strength	Advice Nature (Explicit Recommendation Provided?)	Quality Assessment Standard	n	P_{Accept} (stage 1)	P_{Accept} (stage 2)	Increase (decrease) in P_{Accept} from stage 1 to 2
Weak	No	Absent	15	0.13	0.53	0.40
	No	Present	17	0.18	0.29	0.11
	Yes	Absent	14	0.14	0.36	0.22
	Yes	Present	13	0.15	0.46	0.31
Strong	No	Absent	16	0.44	0.81	0.37
	No	Present	16	0.63	0.75	0.12
	Yes	Absent	18	0.33	0.61	0.28
	Yes	Present	17	0.35	0.24	(0.11)

cant in the absence of QAS ($p = 0.003$) but not in its presence ($p > 0.500$). As shown in Table 3, in the presence of QAS and advice with explicit recommendation, P_{Accept} increases (by 31 percent) from stage 1 to stage 2 when client justification is weak, but decreases (by 11 percent) from stage 1 to stage 2 when client justification is strong. The latter result could be due to auditors' heightened suspicion of client's persuasive tactics (i.e., the provision of a strong justification to influence their decisions), causing them to guard against their own proclivity to increase support for the client-preferred method, particularly given the presence of QAS and explicit recommendation by the technical department. These results potentially explain why P_{Accept} in stage 2 is the lowest in the QAS present, advice with explicit recommendation, and strong client justification condition, and suggest that strong justifications by the client may "backfire" on the client in certain situations.

Overall, the results provide some evidence that auditors strategically use information provided in the advice to bolster their support for the client-preferred method, except for the condition in which the client has provided a strong justification for its preferred method in the presence of QAS and the technical department's advice explicitly recommends the use of the most appropriate method. The finding that, despite the provision of a weak justification by the client for their preferred method, auditors increase their support for the method after receiving the technical department's advice, even in the presence of QAS and an explicit recommendation by the technical department to use the most appropriate method, suggests that auditors are still susceptible to the influence of pro-client directional goals under these conditions.[24]

Further Evidence on the Effects of Explicit Recommendation

Our main hypothesis and results suggest that the provision of an explicit recommendation in the technical department's advice reduces the justifiability of supporting the client-preferred method and, consequently, auditors' propensity to accept the method in the presence of QAS and strong client justifications. To provide supporting evidence for this conclusion, we conducted an additional experiment involving the two advice nature treatments, holding constant the client justification strength (strong) and QAS (present) conditions. The design and instruments used were

[24] Kadous et al. (2003) similarly find that auditors are susceptible to the influence of pro-client directional goal commitment even in the presence of QAS and a weak client justification.

identical to those in the main experiment for the corresponding treatments, with the exception of additional questions inserted in the debriefing questionnaire to elicit additional assessments (discussed below). Participants are 40 audit seniors with an average of 3.5 years of audit experience.[25] Participants in the "advice with explicit recommendation" condition rate the extent to which the technical department explicitly recommends that the client should adopt the most appropriate method to enhance financial reporting quality higher than participants in the advice without explicit recommendation condition (means = 5.5 and 4.6, respectively; $p = 0.043$), indicating a successful manipulation of advice nature.

Consistent with the results of our main experiment, participants are less likely to accept the client-preferred method in stage 2 when the advice contains an explicit recommendation than when it does not ($P_{Accept} = 0.35$ and 0.70, respectively; $p = 0.001$). Their pre-advice decisions in stage 1 do not vary by advice nature ($p = 0.337$). Although only marginally significant, participants rate the client's use of the straight-line method (i.e., the client's preferred method) as less justifiable when the advice contains an explicit recommendation than when it does not (means = 4.3 and 4.9, respectively; $p = 0.074$, one-tailed). We also investigate whether the provision of an explicit recommendation in the technical department's advice influences auditors' decision through its effect on auditors' perceptions of (1) the discretion or flexibility that they have in accepting or rejecting the client-preferred method; (2) the extent to which the audit firm prefers the client to use the most appropriate method; or (3) the usefulness of the technical department's communications in terms of clarifying which accounting method is the most appropriate method or which accounting method the client should use. None of these assessments vary significantly by advice nature (all $p \geq 0.286$).[26] Last, we measure auditors' directional goal commitment (i.e., to build a justifiable case for characterizing the client-preferred method to be an acceptable revenue recognition method in the circumstances), using the five-item scale reported in Kadous et al. (2003, 767), to examine whether it is altered by the presence of an explicit recommendation in the technical department's advice.[27] Results indicate no significant difference in directional goal commitment between the advice nature conditions ($p = 0.347$). Overall, these results support our argument that the provision of an explicit recommendation in the technical department's advice decreases auditors' propensity to accept the client-preferred method by reducing auditors' perceived justifiability of supporting the client-preferred method (in other words, tightening the reasonableness constraints on auditors' motivated reasoning).

V. CONCLUDING REMARKS

This study demonstrates how a key quality control mechanism in audit firms, advice provided by the firm's technical department, interacts with other features in the audit environment (specifically, justifications provided by the client and standards issued by the profession) to jointly influence auditors' propensity to accept client-preferred accounting methods. We find that, in the

[25] The majority (83 percent) has been involved in auditing listed companies, while 45 percent of them have participated in, or prepared for discussions with the clients' audit committees. On average over the past two years, they have consulted the technical department personnel formally (informally) in 30 percent (37 percent), and other colleagues in 68 percent, of their audit engagements. These experience measures do not vary between experimental treatments.

[26] We also elicit participants' agreement with the statement "The technical department's communications on a specific accounting issue reflect the audit firm's preference on the accounting issue" on a seven-point scale (1 = disagree completely; 7 = agree completely). Participants generally agree with this statement (mean = 4.96) regardless of advice nature ($p > 0.500$). Consistent with the results of the main experiment, participants also agree with the statement "When there is more than one method of accounting for a transaction that could be deemed acceptable under GAAP, auditors should require the client to use the most appropriate accounting method" (mean = 5.8) regardless of advice nature ($p > 0.500$).

[27] Results of factor analysis on responses to the five items indicate only one factor with an eigenvalue greater than 1 (3.08). This factor explains 62 percent of the variance in the measures and Cronbach alpha is 0.839.

absence of any advice, auditors' propensity to support a client-preferred method is significantly influenced by the strength of the client's justification, but not by the presence of an SAS No. 90-like quality assessment standard. In contrast, when auditors are provided with advice from the firm's technical department indicating that an alternative method is the most appropriate method, auditors' propensity to accept the client-preferred method is reduced by the presence of the quality assessment standard, but only when the advice explicitly recommends the use of the most appropriate method and the client has provided a relatively strong justification for its preferred method.

Our findings highlight the role and limitations of in-house consultation and advice in enhancing auditors' decisions, and suggest that the nature of advice sought by, and presented to, auditors matters. Specifically, advice that merely reduces the ambiguity surrounding the relative quality of alternative accounting methods may not be sufficient to counteract auditors' proclivity to support a less appropriate accounting method preferred by the client, even in the presence of a quality assessment standard. Including an explicit recommendation in the advice to use the most appropriate method is instrumental in helping auditors overcome the influence of strong client justifications, but only in the presence of a quality assessment standard. These findings demonstrate the complementary roles that professional standards and audit firms' technical departments play in enhancing auditors' accounting policy decisions.

Our within-subject analyses of changes in auditors' decisions between stage 1 and stage 2 of the experiment indicate that, except for the condition in which the client has provided a strong justification for its preferred method in the presence of QAS and the technical department's advice explicitly recommends the use of the most appropriate method, auditors strategically use information in the advice (that the client-preferred method is an acceptable, albeit not the most appropriate, method under GAAP) to bolster their support for the client-preferred method even when it may not be the most appropriate method. We find that, except for the above-mentioned condition, the proportion of auditors who accept the client-preferred method actually increases (rather than decreases) in the presence of the advice. These results suggest that the presence of advice could lead to an unintended consequence of bolstering auditors' support for client-preferred methods, especially if the advice suggests that these methods, while not the most appropriate, do not lead to an outright violation of GAAP.

Several limitations of our study should be noted. First, we examine the effects of advice received from a single source (i.e., the audit firm's technical department). However, there are different sources of advice within the firm with which auditors can potentially consult, and our results may not extend to other sources of advice. Future research could examine whether differential attributes among alternative sources of advice (such as expertise) influence auditors' propensity to seek and use advice from specific sources (e.g., Arnold and O'Connor 1999).

Second, prior literature suggests that individuals are more likely to follow solicited advice relative to unsolicited advice (Gibbons 2003). Although our participants indicate that they are likely to consult the firm's technical department on the accounting issue (see footnote 19), it is unclear whether our results would hold had we allowed them to solicit advice themselves rather than providing them with unsolicited advice. Future research could examine whether our results are sensitive to whether advice is solicited.

Third, we examine a context in which the advice provides some support for the client-preferred method by indicating that it is acceptable under GAAP, albeit not the most appropriate method. While this context is directly relevant for testing the efficacy of a quality assessment standard, which requires auditors to assess the appropriateness, not just acceptability, of client's accounting methods, it also provides an opportunity for auditors to use the advice to bolster their support for the client-preferred method. Thus, our results may not hold had the advice indicated that the client-preferred method is of questionable merit, as auditors would find it less justifiable to support such a method even in the absence of a quality assessment standard or an explicit

recommendation. Nonetheless, future research could investigate whether auditors continue to support the client's preferred method even in the presence of advice that indicates that the client's method is of questionable merit. Another avenue for future research is to examine how the availability of potentially conflicting advice from multiple sources influences auditors' decisions (e.g., Sniezek and Buckley 1995; Yaniv and Milyavsky 2007).

Fourth, we examine factors influencing the decisions of auditors as advice recipients, not advice-givers. Thus, our study does not address the question of whether and how various contextual features (such as client preferences) influence the quality of advice provided to auditors by advice-givers such as the technical department personnel. To the extent that client preferences influence the quality of advice provided by advice-givers, audit effectiveness could be compromised.

Fifth, while our participants are experienced audit seniors and managers, our results might not apply to audit partners who make the ultimate decisions on clients' accounting methods. Nonetheless, understanding the effects of advice on the judgments and decisions of audit seniors and managers is important, given that audit seniors and managers are involved in consultations with the technical department and other colleagues and regularly provide inputs to audit partners in the resolution of clients' accounting issues. Future research could examine whether our results vary depending on the auditors' rank (i.e., partners versus managers/seniors) or other individual auditor attributes that our study does not capture, such as tacit managerial knowledge (e.g., Shankar and Tan 2006). Future research could also investigate whether other client features such as audit committee effectiveness (e.g., Ng and Tan 2003) or additional regulatory measures introduced under the Sarbanes-Oxley Act of 2002 (e.g., Nelson and Tan 2005) moderate the findings reported here.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 1999. *Audit Committee Communications*. Statement of Auditing Standards No. 90. New York, NY: AICPA.

———. 2006. *Conceptual Framework for AICPA Independence Standards*. ET Section 100.01. New York, NY: AICPA.

———. 2009. *A Firm's System of Quality Control*. Statement on Quality Control Standards No. 7. New York, NY: AICPA.

Anderson, U., K. Kadous, and L. Koonce. 2004. The role of incentives to manage earnings and quantification in auditors' evaluation of management-provided information. *Auditing: A Journal of Practice & Theory* 23 (March): 11–27.

Arnold, J. A., and K. M. O'Connor. 1999. Ombudsperson or peers? The effect of third-party expertise and recommendations on negotiation. *The Journal of Applied Psychology* 84: 776–785.

Bame-Aldred, C. W., and T. Kida. 2007. A comparison of auditor and client initial negotiation positions and tactics. *Accounting, Organizations and Society* 32: 497–511.

Bonaccio, S., and R. S. Dalal. 2006. Advice taking and decision-making: An integrative review of the literature and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes* 101: 127–151.

Danos, P., J. W. Eichenseher, and D. L. Holt. 1989. Specialized knowledge and its communication in auditing. *Contemporary Accounting Research* 6: 91–109.

Gibbins, M., and C. Emby. 1985. Evidence on the nature of professional judgment in public accounting. Auditing Research Symposium (1984), 181–212, Office of Accounting Research, University of Illinois at Urbana–Champaign.

Gibbons, A. A. M. 2003. Alternative forms of advice in natural decision settings. Master's thesis, University of Illinois at Urbana–Champaign.

Hackenbrack, K., and M. Nelson. 1996. Auditors' incentives and their application of financial accounting standards. *The Accounting Review* 71 (January): 43–59.

Harvey, N., and I. Fischer. 1997. Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes* 70: 117–133.

Hsee, C. K. 1995. Elastic justification: How tempting but task-irrelevant factors influence decisions. *Organizational Behavior and Human Decision Processes* 62: 330–337.

International Auditing and Assurance Standards Board (IAASB). 2009a. *Communication with Those Charged with Governance.* International Standard on Auditing 260. New York, NY: International Federation of Accountants.

———. 2009b. *International Standard on Quality Control 1.* New York, NY: International Federation of Accountants.

Jenkins, J. G., and C. M. Haynes. 2003. The persuasiveness of client preferences: An investigation of the impact of preference timing and client credibility. *Auditing: A Journal of Practice & Theory* 22 (March): 144–154.

Kadous, K., J. Kennedy, and M. E. Peecher. 2003. The effect of quality assessment and directional goal commitment on auditors' assessment of client-preferred accounting methods. *The Accounting Review* 78 (July): 759–778.

Kennedy, J., D. N. Kleinmuntz, and M. E. Peecher. 1997. Determinants of the justifiability of performance in ill-structured tasks. *Journal of Accounting Research* 35 (Supplement): 105–123.

Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108: 480–498.

Libby, R., and H.-T. Tan. 1999. Analysts' reactions to warnings of negative earnings surprises. *Journal of Accounting Research* 37 (Autumn): 415–435.

Nelson, M., and H.-T. Tan. 2005. Judgment and decision making research in auditing: A task, person, and interpersonal interaction perspective. *Auditing: A Journal of Practice & Theory* 24 (Supplement): 41–71.

Ng, T. B.-P., and H.-T. Tan. 2003. Effects of authoritative guidance availability and audit committee effectiveness on auditors' judgments in an auditor-client negotiation context. *The Accounting Review* 78 (July): 801–818.

Perkins, J. D. 2003. Informal consultation in public accounting: A strategic view and an experimental investigation. Ph.D. thesis, University of Illinois at Urbana–Champaign.

Petty, R. E., and J. T. Cacioppo. 1986. The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology* 19: 124–205.

Salterio, S. 1996. The effects of precedents and client position on auditors' financial accounting policy judgments. *Accounting, Organizations and Society* 21: 467–486.

———, and R. Denham. 1997. Accounting consultation units: An organizational memory analysis. *Contemporary Accounting Research* 14: 669–691.

———, and L. Koonce. 1997. The persuasiveness of audit evidence: The case of accounting policy decisions. *Accounting, Organizations and Society* 22: 573–587.

Shafer, W. E., A. A. Ketchand, and R. E. Morris. 2004. Auditors' willingness to advocate client-preferred accounting principles. *Journal of Business Ethics* 52: 213–227.

Sniezek, J. A., and T. Buckley. 1995. Cueing and cognitive conflict in judge-advisor decision making. *Organizational Behavior and Human Decision Processes* 62: 159–174.

Shankar, P. G., and H.-T. Tan. 2006. Determinants of audit preparers' workpaper justifications. *The Accounting Review* 81: 473–495.

Trompeter, G. 1994. The effect of partner compensation schemes and GAAP on audit partner judgments. *Auditing: A Journal of Practice & Theory* 13 (Fall): 56–68.

Yaniv, I. 2004. Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes* 93: 1–13.

———, and E. Kleinberger. 2000. Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes* 83: 260–281.

———, and M. Milyavsky. 2007. Using advice from multiple sources to revise and improve judgments. *Organizational Behavior and Human Decision Processes* 103: 107–120.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1763–1786

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1763

The Effects of a Supervisor's Active Intervention in Subordinates' Judgments, Directional Goals, and Perceived Technical Knowledge Advantage on Audit Team Judgments

Mark E. Peecher
University of Illinois at Urbana–Champaign

M. David Piercey
University of Massachusetts Amherst

Jay S. Rich
Illinois State University

Richard M. Tubbs
The University of Iowa

ABSTRACT: Prior research shows that an audit supervisor's active intervention in a subordinate's judgment distorts that judgment. However, subordinates' judgments are only one input into audit team judgments. How do supervisors finalize audit team judgments after actively intervening in their subordinates' judgments? In an experiment using audit teams, supervisors with weaker or stronger goals to reach a client-preferred conclusion either were or were not asked to first actively coach a subordinate's judgment (i.e., active intervention) before reviewing it and finalizing the audit team's judgment. Supervisors' intervention influenced subordinates' inputs, which, in turn, supervisors incorporated into their final judgments. More interestingly, intervention biased supervisors' final judgments, controlling for supervisor directional goal strength and for concurrent effects on subordinates' inputs. However, supervisors distorted their judg-

We thank two anonymous reviewers, S. Kachelmeier (editor), R. Bloomfield, J. Brazel, T. Kida, R. Libby, M. Nelson, E. O'Donnell, B. Pomeroy, J. Smith, H.-T. Tan, workshop participants at Cornell University, Loyola University Chicago, The Ohio State University, as well as conference participants at the 2008 ISAR Conference for their helpful comments. An earlier paper based only on Experiment 2 also benefited from the helpful comments of R. Balakrishnan, M. Bamber, J. Herbold, K. Kadous, D. Kleinmuntz, S. Krische, K. Krumwiede, R. Ramsay, H.-T. Tan, J. Wilks, M. Zimbelman, conference participants at the AAA Audit Midyear Meeting, and workshop participants at Brigham Young University, Florida State University, The Ohio State University, and The University of Georgia. We thank K. Fanning for research assistance. Finally, we thank the audit firm and professionals who participated.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: July 2008
Accepted: February 2010
Published Online: August 2010

ments less as they perceived a larger technical knowledge advantage over subordinates. In a second experiment, auditors appear aware of the bias-reducing knowledge advantage effects but unaware of the bias-increasing active intervention effects. We discuss implications for audit team judgments and audit quality control.

Keywords: *audit team judgments; client pressure; directional goals; active intervention; technical knowledge.*

Data Availability: *Data are available from the authors.*

I. INTRODUCTION

Audit supervisors ideally guide and review subordinate auditors' work in ways that improve the judgments of subordinates and, ultimately, of audit teams (e.g., Bamber et al. 1988; Gibbins and Trotman 2002; Phillips 1999; Rich 2004; Rich et al. 1997a; Shankar and Tan 2006; Tan and Tan 2008). Prior research shows that supervisors who allude to preferred audit conclusions when "coaching" or actively intervening in subordinates' judgments can bias the subordinates' judgments (e.g., Peecher 1996; Brown et al. 1999; Wilks 2002; Piercey 2009). The conditions under which subordinate judgment bias distorts audit team judgments, however, are unclear, as subordinates' judgments are only an input. Supervisors sometimes form their own initial judgments, and they inevitably finalize judgments for the audit team (Rich et al. 1997a).

This observation raises the research question of how supervisors who have directional goals and guide their subordinates via active intervention finalize audit team judgments (as compared to supervisors who have the same directional goals, but do not intervene). We examine this question by investigating the effects of a supervisor's active intervention, directional goals, and perceived technical knowledge advantage over subordinates on the supervisor's finalized judgments for the audit team. We also examine whether auditors understand how these factors moderate bias in audit team judgments.

That supervisors with directional goals commonly use active intervention to coach subordinates' judgments raises two worrisome possibilities for audit team judgments. One is that, after coaching the subordinate, supervisors would anchor on the subordinate's judgment, preserving any intervention-induced distortions in the audit team's judgment. This possibility would validate concerns raised elsewhere (e.g., Peecher 1996; Wilks 2002). An even more worrisome possibility is that active intervention incrementally biases the supervisors' judgments, controlling for concurrent effects of intervention on subordinates' judgments. If so, active intervention would likely pose greater audit effectiveness concerns than previously understood from prior theory. In the auditing environment, client pressures can create directional goals for auditors to reach biased judgments of financial statements. Research suggests that such directional goals to reach client-preferred auditing judgments contributed to many prominent accounting scandals, prompted reforms in the Sarbanes-Oxley Act (SOX), and continue to put audit effectiveness at risk, even post-SOX (Moore et al. 2006; Nelson 2006; Bazerman et al. 2006).

To investigate our research questions, we conduct two experiments. In the first experiment, participants were industry-specialist auditors in supervisor-subordinate teams. The audit teams' task was to assess the discount rate used to value a financial asset securitization, which provides a relevant and timely setting for testing our hypotheses.[1] Supervisors had either stronger or weaker directional goals to reach client-preferred accounting estimates and either were or were not asked

[1] Empirical evidence suggests managers use the discount rate for securitizations to manage earnings (Dechow and Shakespeare 2009). For example, regarding the audited financial statements of New Century Financial, Harris (2008, 1) quotes bankruptcy manager Michael Missal's allegation that the auditors "'acquiesced to a lot of what New Century wanted to do' ... One problem was 'the discount rate being used to compute the residual interest valuation ... The

to actively coach a subordinate's preliminary audit judgment before reviewing it and finalizing the audit team's judgment. We find that subordinates' judgments (influenced by the active intervention) are reflected in supervisors' finalized judgment. Further, intervention leads to incremental bias in supervisors' finalized judgment beyond that attributable to any effects on subordinates' judgments.

We also examine how supervisors' perceptions of their own technical knowledge (compared to that of their subordinates) influences audit-team judgment bias. We hypothesize that this advantage would decrease such bias, based on the idea that supervisors sense less of a need to persuade themselves about the reasonableness of their preferred audit conclusions when they believe themselves to have greater subject-matter expertise within the audit team. As predicted, a perceived technical knowledge advantage reduces the biasing effects of supervisors' directional goals on finalized audit team judgments. Thus, while coaching via active intervention exacerbates supervisors' goal-consistent bias, there are conditions under which supervisors' preferences are less apt to cause audit team judgment bias.

In a second experiment, we examine auditors' perceptions of how these factors influence audit-team judgment bias. This experiment is motivated by the idea that auditors must first understand determinants of audit-team judgment bias to be able to design quality control policies that effectively mitigate it (Lerner and Tetlock 1999; Moore and Loewenstein 2004). We find that, while auditors understand the bias-*reducing* effects of a supervisor's knowledge advantage over a subordinate, they do not understand the bias-*inflating* effects of supervisors' coaching via active intervention. Quality control activities would likely improve if auditors' understanding were improved.

Our theory and experimental findings have potentially important audit effectiveness implications. As Kadous et al. (2008) note, motivated reasoning occurs when accountants change their judgments to become more consistent with external client preferences. We posit active intervention increases the tendency of audit supervisors to do exactly that. As Wilks (2002) points out, intervening earlier in a subordinate's judgment process could undermine audit quality because of distortionary effects on subordinates' judgments. Our findings suggest that, by intervening, supervisors not only influence their subordinates, but also distort their *own* judgments. When supervisors have directional goals to reach client-preferred conclusions, our study suggests that their judgments would be less biased by client preferences if supervisors did not actively coach their subordinates and instead received more independent judgments from their subordinates.

Audit firms, standard-setters, and regulators could consider training and quality control policies that take better account of these sources of audit-team judgment bias. As one example, individual auditors would likely benefit from greater awareness of how supervisors' active intervention can elevate audit-team judgment bias, especially when conditions suggest the supervisor also has a client-preferred directional goal. As another example, the PCAOB (2008, 1) has identified *cold* or *concurring-partner* reviews (i.e., back-end reviews by audit partners without intervening in the audit as it unfolds) as "probably the most" frequent audit deficiency in current practice, and the SEC (2010) recently approved the PCAOB's AS No. 7, a new auditing standard that increases their importance. Our study suggests that such reviews would be less vulnerable to this new source of bias, as they usually are done without intervention. Our findings potentially also inform other organizational contexts in which supervisors can coach a subordinate's input and have directional pressures to reach preferred conclusions.

---

company wanted it as low as possible,' so that the valuations could be as high as possible." Such subjective estimates are common in financial statement audits (Abdolmohammadi 1999) and provide opportunity for directional goals to influence supervisors' judgments (Kadous et al. 2003).

We develop a theoretical framework and hypotheses in Section II, describe the method and results of Experiments 1 and 2 in Sections III–VI, and conclude in Section VII.

II. THEORY AND HYPOTHESES

Supervisors' Directional Goals and Active Intervention in Subordinates' Judgments

As Kunda (1990, 480) notes, "motivated reasoning phenomena ... fall into two major categories: those in which the motive is to arrive at an accurate conclusion ... and those in which the motive is to arrive at a particular, directional conclusion." She uses the terms *accuracy goals* and *directional goals* to distinguish these broad categories. Directional goals can self-originate in a person, such as when investors' preferred investment outcomes depend on whether they hold a short or long position (Hales 2007), or a person can inherit directional goals from socially significant others, such as when auditors inherit directional goals from their supervisors (Piercey 2009) or from their clients (Kadous et al. 2003).[2]

Regardless of their origination, directional goals' activation and influence need not be, and usually are not, fully consciously processed as if they were intentional goals (Chartrand and Bargh 1996; Moore et al. 2006, 18; Wilks 2002, 54). Indeed, at the conscious level, decision makers who have directional goals try to convince themselves and others that they are reasonable and impartial, but this is only an "illusion of objectivity," as these goals significantly bias their judgments (Pyszczynski and Greenberg 1987, 302).

The need to maintain reasonableness limits how much bias directional goals can infuse into a person's judgments (Kunda 1990). However, factors designed to encourage auditors to think more deeply and reasonably do not necessarily reduce and, in fact, can increase the amount of bias caused by underlying directional goals (cf. Kunda 1990, 491). In Kadous et al. (2003), for example, auditors inherit a directional goal to accept an aggressive, client-preferred accounting method. About half the auditors are asked to identify the most appropriate accounting method and to provide a written justification for their choice before deciding whether to accept the client's method. Compared to auditors not asked to do these tasks, these auditors became significantly more likely to accept the client-preferred accounting method. This result occurred even though most auditors (70 of 98) characterized the client-preferred method as conceptually suboptimal when identifying the most appropriate method.

An important contextual factor that likely moderates how deeply supervisors with directional goals process information during the review process is whether they actively intervene in or "coach" their subordinate's judgment. Related prior research indicates that when subordinates inherit client-preferred directional goals from supervisors, subordinates change their judgment processes to favor client-preferred audit conclusions (e.g., Brown et al. 1999; Peecher 1996; Piercey 2009; Shankar and Tan 2006; Turner 2001; Wilks 2002). The supervisors in these studies' experiments are hypothetical, so these studies cannot empirically examine whether supervisors with directional goals maintain objectivity in the nature of messages they convey to subordinates. These studies also cannot address the arguably more important question of whether active intervention moderates the influence of supervisors' directional goals on their own finalized judgments for audit teams.

Such moderating effects are potentially critical because supervisors commonly engage in active intervention using a variety of modes. One mode is "real-time" review, which began in the mid-1990s as a means to supplement and, for lower-risk conditions, even supplant traditional "back-end" review (e.g., Rich et al. 1997b; Winograd et al. 2000; Gibbins and Trotman 2002).

[2] For other examples of accountants inheriting directional goals from their clients and/or supervisors see, e.g., Brown et al. 1999, Kadous and Magro 2001, Kadous et al. 2008, Peecher 1996, Salterio and Koonce 1997, and Wilks 2002.

While the prevalence of real-time review appears to have waned recently, many modes of active intervention continue to flourish. Common modes include informal or formal face-to-face discussions, interviews, and various kinds of electronic communications (e.g., Agoglia et al. 2008; Brazel et al. 2004; Shankar and Tan 2006; Tan and Tan 2008). Recent field studies report that supervisors use discussions, interviews, and electronic communications, and modes of active intervention likely will continue to evolve (Fargher et al. 2005; Payne et al. 2010; Vera-Muñoz et al. 2006).

Regardless of communication mode, our active intervention construct refers to whether and the degree to which a supervisor "coaches" or "guides" a subordinate's judgment before reviewing it as opposed to not coaching or guiding and reviewing a more independent judgment from a subordinate. Active intervention includes, to varying degrees, intervention cognition and intervention content. Intervention cognition refers to the automatic and controlled cognition that occurs throughout intervention.[3] It is goal-directed, including supervisors' private thinking and their thinking about what (and what not) to convey to the subordinate and why as well as the best way to say things (cf. stylization in Rich et al. [1997a] and Rich [2004]). Intervention content refers to properties of messages communicated to subordinates (e.g., one- versus two-sided, include versus omit preferred conclusions, start versus finish with strongest arguments, etc.; Petty and Wegener 1998). Whenever supervisors convey intervention content to subordinates, some incremental degree of intervention cognition has occurred.

Motivated reasoning theory holds that even individuals with strong directional goals strive to reach conclusions that seem defensible to themselves and impartial to others (Kunda 1990). Thus, they take steps to maintain an illusion of objectivity (Pyszczynski and Greenberg 1987). As Kunda (1990, 483) notes:

> [T]he objectivity of this justification construction process is illusory because *people do not realize* that the process is biased by their goals, that they are accessing only a subset of their relevant knowledge, that they would probably access different beliefs and rules in the presence of different directional goals, and that they might even be capable of justifying different conclusions on different occasions. (emphasis added)

One way for audit supervisors who actively intervene in their subordinates' judgment to maintain this illusion would be to edit or censor what they say to subordinates, trying to avoid bias in intervention content. Overt bias in communicated content would impair justifiability. Nevertheless, their own, deeper cognition in thinking about what to and not to convey includes automatic goal-driven judgment processes that bias supervisors toward reaching and justifying preferred conclusions (cf. Kadous et al. 2003; Kunda 1990; Piercey 2009). Therefore, the attempt to provide bias-free content to subordinates, even if successful, likely does not translate into bias-free cognition and so gives supervisors with directional goals a false sense of security about the justifiability (i.e., neutrality or objectivity) of their conclusions.

Because active intervention triggers more cognition directed at self-justifying supervisors' preferred judgments, actively intervening supervisors likely reach audit judgments that are more biased by their directional goals than those who have the *same* goals but do not intervene, controlling for any concurrent effects of intervention content on their subordinate's judgments. This reasoning supports the following hypothesis:[4]

---

[3] Automatic cognition, in contrast to controlled cognition, occurs without awareness and is hard to stop once initiated (Sherman et al. 2008).

[4] We do not posit a hypothesis regarding the influence of supervisors' directional goals on the intervention content since motivated reasoning theory would allow for a null (no-effects) prediction. We do, however, measure, report, and control for the influence of supervisors' intervention content on the subordinate's judgment as part of testing H1.

H1: A supervisor's intervention increases the effects of directional goals to reach a preferred conclusion on his/her finalized judgment for the audit team, even after controlling for any effects of intervention on the subordinate's initial judgment.

Note that one cannot attribute empirical support for H1 solely to supervisors using their role to adjust for a subordinate not reaching the supervisor's preferred conclusion, because non-intervening supervisors with the same directional goals can make the same adjustment. In that case, the finalized judgments of supervisors who intervene would equal those of supervisors who do not intervene but have the same goals. However, if judgments of intervening supervisors differ from those who do not intervene, while holding constant their directional goals and controlling for subordinates' judgments (H1), active intervention must create additional distortion in supervisors' judgments, via its cognition component.

Supervisors' Directional Goals and Perceived Technical Knowledge Advantage over Subordinates

Absent theory, one could imagine that a perceived technical knowledge advantage over subordinates could increase or decrease supervisor judgment bias. As one possibility, supervisors could consciously or unconsciously lever such a perceived advantage to instill more directional-goal consistent bias into the audit team's conclusions. Alternatively, such an advantage could prompt supervisors to be objective, better fulfilling their quality-control role (cf. Libby and Trotman 1993). Motivated reasoning theory, interestingly, supports the alternate possibility, for two reasons.

First, motivated reasoning theory holds that individuals with directional goals limit the amount they bias their judgments, as noted above, to maintain the illusion of objectivity to themselves and others. This need to appear objective creates *reasonableness constraints* on motivated reasoning (see Kunda 1990). Factors that increase the salience of reasonableness constraints curtail the extent to which supervisors with directional goals instill bias into audit team conclusions (e.g., Cuccia et al. 1995; Salterio and Koonce 1997; Piercey 2009; Ng and Shankar 2010). Supervisors who perceive a larger technical knowledge advantage over subordinates likely are, on average, more sensitive to reasonableness constraints and, therefore, less likely to bias their judgments. For example, supervisors who feel that they provide the majority of the technical knowledge within the audit team for assessing discount rates are likely to become more sensitive to the reasonableness of a lenient discount rate for the valuation of a fairly risky asset when finalizing team judgments.

Second, motivated reasoning theory also holds that individuals' directional reasoning is *instrumental*. That is, they engage in *just enough* one-side reasoning to attain their goal, even if more bias would still be viewed as reasonable (Boiney et al. 1997). As supervisors feel that their technical knowledge dominates their subordinates', they likely sense less of a need to self-justify their judgments and therefore engage in less directional reasoning. Consequently, their intervention cognition entails less rehearsal of various arguments to justify reaching a preference-consistent conclusion. Thus, motivated reasoning theory suggests the following hypothesis:

H2: A supervisor's perceived technical knowledge advantage (over subordinates) reduces the effects of a supervisor's directional goals on his (or her) finalized assessments for the audit team.[5]

[5] Theory does not warrant predicting that the effects of H2 will occur differently in the intervention and no-intervention conditions (i.e., a three-way interaction). In our development of H1, we posit that intervention affects auditors' judg-

III. EXPERIMENT 1: METHOD

Participants

We recruited 61 financial services industry specialist auditors to participate in a valuation task in a financial services setting. The vast majority of auditors (95 percent) indicated that they had financial services industry expertise. The auditors also provided experience information in the post-test questionnaire, summarized in Table 1.

We assigned auditors to the role of supervisor or subordinate based on experience information provided by their firm, so we had a potential of 30 supervisor-subordinate pairs.[6] Nearly 90 percent of the subordinate participants were seniors or staff/assistants. Of the supervisor participants, 55 percent were in-charges and 45 percent were seniors. Compared to subordinates, supervisors had more experience in professional accounting ($t = 6.9$, $p < 0.001$), and in reviewing

TABLE 1

Experiment 1: Audit Teams

	Subordinates	Supervisors
Financial Services Industry Expertise	28 (97%)	27 (93%)
Rank		
Assistant/First-Year	2 (7%)	0
Staff	10 (34%)	0
Senior	14 (48%)	13 (45%)
In-Charge	3 (10%)	16 (55%)
Experience in Professional Accounting		
Mean	1.8 years	4.1 years
Median	1.8 years	4.0 years
Minimum	0.0 years	1.5 years
Maximum	4.0 years	8.3 years
Standard Deviation	0.9 years	1.5 years
Experience as a Supervisor		
Mean	0.4 years	2.2 years
Median	0.3 years	2.0 years
Minimum	0.0 years	0.2 years
Maximum	1.5 years	5.0 years
Standard Deviation	0.4 years	1.2 years
n	29	29

ments in unconscious ways (results of Experiments 1 and 2 support this). If we expected that auditors are aware of intervention effects, we might expect supervisors who perceive a larger knowledge advantage to attempt to offset this effect. However, we do not believe auditors are conscious of intervention effects and therefore do not predict this. Since auditors *do* believe goals to bias their judgments (as in Wilks' [2002] Experiment 2 and our own Experiment 2), we do expect supervisors who perceive a larger knowledge advantage to try to offset their own directional goals, as H2 predicts.

[6] The sample excludes the 61st auditor, with whom we paired one of the authors present at the experiment.

subordinates' work ($t = 6.8$, $p < 0.001$).[7] Each auditor knew his/her role, but not the identity of his/her counterpart, who was in a different room. We used this control to prevent unmeasured peripheral social cues from influencing our results.

We excluded one outlying observation (the supervisor's finalized judgment for the audit team was an unreasonable 50 percent discount rate) to avoid observing spurious effects.[8] This provided us with a sample of 29 usable supervisor-subordinate teams.

Design

Figure 1 summarizes the experimental design, and Table 2 summarizes variables in the order that participants encountered them in the task. All participants read a case about a hypothetical auditing client and information relating to the client's valuation of a financial asset. The dependent variable of interest is each supervisor's finalized judgment (for the audit team) of a discount rate for the valuation of a client's asset (*FinalJgmt*). Lower values indicate more aggressive (i.e., client-preferred) valuations. All supervisors received input (a preliminary discount rate) from a subordinate (*SubJgmt*) before forming the finalized judgment (*FinalJgmt*). However, half of the supervisors were first asked to provide guidance for the subordinate's preliminary judgment. Audit teams were randomly assigned to this intervention (*I*) treatment, the experiment's only manipulated factor (Figure 1).

Experimental Task

Introduction

Both supervisors and subordinates read an introduction that described a hypothetical auditing client, Retail Credit, Inc. (RCI), the financing arm of a clothing retailer. The supervisors' instrument stated that they would receive input from another professional from their firm, whom they should treat as a subordinate under their supervision. Similarly, the subordinates' instrument stated that another professional, whom they should treat as a supervisor, would review their work.

Directional Goals to Reach an Engagement Partner's Preferred Conclusion

Supervisors inherited a directional goal from their partner for the audit team. In particular, the supervisors' instruments described the engagement partner as generally preferring to allow client-preferred positions on subjective accounting estimates and included a question that asked participants to interpret the partner's preferences for the audit team judgments. Supervisors used an 11-point scale to indicate whether they thought the partner preferred more conservative/income-decreasing estimates (−5), neutral estimates (0), or less conservative/income-increasing estimates

[7] As part of the experimental control that kept audit team members anonymous to one another (as mentioned above), no participants knew how many years of experience the other auditor with whom they were paired had. Additionally, as explained in the description of the task (below) the instrument simply asked participants to assume for the purposes of the case their role as supervisor or subordinate, and to assume the opposite role for their anonymous team member. Thus, *a priori*, the anonymous teams should prevent inadvertent differences in self-reported actual experience (or other unmeasured social cues) between supervisors and subordinates (Table 1) from influencing our results. Consistent with this, we do not detect any effects of actual experience gap in our results *ex post*.

[8] All other observations ranged from 9 percent to 17 percent. The outlying observation responded with a 50 percent discount rate despite receiving case information of 8 percent to 14 percent as a range for valuations of instruments with "typical" risk in the industry and a subordinate's input of 11 percent. Multiple diagnostic tests suggest that this observation is statistically outlying as well as contextually nonsensical (Neter et al. 1996).

FIGURE 1
Experiment 1: Design

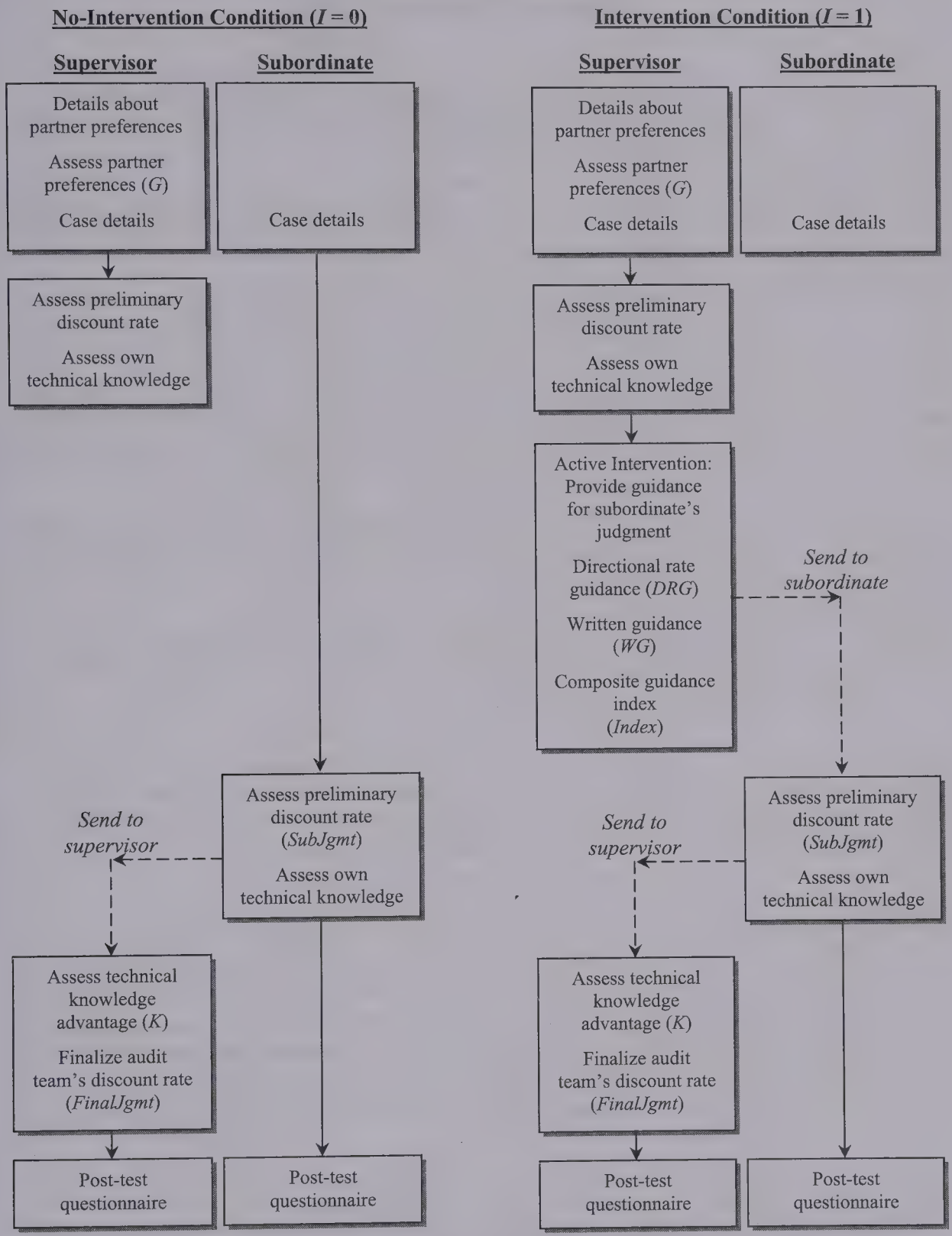


TABLE 2

Experiment 1
Variables in Chronological Order

Variable		Description
g	=	The supervisor's assessment of partner's preferences for allowing more client-preferred (+5), neutral (0), or conservative (−5) valuations of the illiquid financial asset. Higher values indicate stronger directional goals for client-preferred valuations.
G	=	Supervisors' mean-centered goals (i.e., $G = g - \bar{g}$).
I	=	Manipulated intervention. Each supervisor either does (coded 1) or does not (coded 0) have an opportunity to provide directional rate guidance (*DRG*) or written guidance (*WG*) to his or her subordinates before the subordinate provides a preliminary discount rate (*SubJgmt*).
DRG	=	The directional rate guidance that supervisors who intervene provide to their subordinates, indicating the extent to which the discount rate should be above (+5), at (0), or below (−5) industry averages. Smaller values indicate more aggressive, client-preferred valuations.
WG	=	The written guidance that supervisors who intervene provide to subordinates, indicating whether the discount rate should be higher (positive) or lower (negative). *WG* is computed as the number of written points supporting a higher discount rate (WG_{high}) minus the number of written points supporting a lower discount rate (WG_{low}; i.e., $WG = WG_{high} - WG_{low}$).
Index	=	A composite index that summarizes the contents of supervisors' intervention (*DRG* and *WG*), indicating whether each supervisor's guidance collectively suggests higher or lower discount rates to his or her subordinate. Index is the first principal component from a factor analysis of *DRG*, WG_{high}, and WG_{low}.
SubJgmt	=	The subordinate's judgment of a discount rate, to be reviewed as a preliminary input by his or her supervisor, who will finalize a judgment (*FinalJgmt*) for the audit team.
k	=	The supervisor's perceived technical knowledge advantage (+5) or disadvantage (−5), relative to the subordinate.
K	=	Supervisors' mean-centered technical knowledge advantage (i.e., $K = k - \bar{k}$).
FinalJgmt	=	The supervisor's finalized judgment of a discount rate for the audit team. Smaller values indicate more aggressive, client-preferred valuations.

(+5).[9]

We define the variable g as supervisors' perceptions of the partner's preferences. A well-replicated finding is that variation in an engagement partner's preferences affects auditors' directional goals (g) to reach more conservative or more client-preferred conclusions (e.g., Wilks 2002; Brown et al. 1999; Peecher 1996). For interpretive reasons discussed in the results (footnote 15), we also define G as the mean-centered transformation of g (i.e., $G = g - \bar{g}$). We treat larger values of g and G as indicators of stronger supervisor directional goals of favoring client-preferred, income-increasing estimates.

[9] We fashioned the language communicating the partner's preference after the *optimistic* condition in Piercey (2009). Overall, participants perceived the partner's preferences for their judgments to favor income-increasing estimates ($t = 3.76$, $p < 0.001$). As a realism check, a post-test question also asked auditors about the degree of their awareness (in practice) of supervisors' preferences regarding the latitude that partners allow management in making valuation estimates. On a 0 "never aware" to 10 "extremely aware" scale, mean ratings are 6.4 for subordinates and 6.2 for supervisors. Both groups' ratings are above 0 (p's < 0.001) and the midpoint of "somewhat aware" (p's < 0.003).

The subordinates' instrument was silent about partner and manager preferences. We omitted such information to reinforce our intervention (*I*) manipulation: supervisors' intervention content would be the only way that subordinates could learn about such matters and that depended on whether they were randomly assigned to the intervention or no-intervention condition.[10]

Case Details

Supervisors and subordinates then read identical information about the valuation of a financial asset.[11] The client had securitized its credit card accounts receivables (selling them while retaining a 10 percent interest in their cash flows). The instrument described how factors such as the risk-free rate, default risk, credit ratings, collection terms, first-loss terms, pre-payment rates, macroeconomic factors, liquidity, and volatility could affect the discount rate used to value this retained interest. It provided 8 percent to 14 percent as a typical range of discount rates used by credit card issuers to value retained interest, and 9 percent as the discount rate used by MBNA, the largest and one of the most stable credit card issuers. The provision of typical discount rates for comparison makes this an analytical procedures task (Ricchiute 2006).

The instrument also stressed that (1) *lower* discount rates *increase* the fair value estimate, which, (2) in turn, *increases* the client's realized gain on securitization. It then asked two questions to quiz participants on those relationships. The instrument then provided supervisors and subordinates identical details about RCI's portfolio of retained interest, some (other) details suggesting its risk was above (below) the industry average.

Supervisor Intervention and the Subordinate's Preliminary Judgment

As our research question focuses on how active intervention moderates the influence of supervisors' directional goals on their finalized judgments, we desired a strong manipulation of active intervention.[12] So, in one level, no intervention occurred at all, while we prompted both intervention content and cognition in the other level. In the no-intervention condition, supervisors first assessed a preliminary discount rate and their own task-relevant technical knowledge, and then received the subordinates' independently-formed discount rate judgments (*SubJgmt*), along with the subordinates' self-assessed technical knowledge.

In the active intervention condition, supervisors also first assessed a preliminary discount rate and their own task-relevant knowledge. However, their case materials then advised them that their subordinate expected to receive some guidance from them before coming up with a preliminary judgment for the discount rate (*SubJgmt*). These supervisors were asked to rate the importance of four case details (e.g., the credit ratings of RCI's customers) on a 1 to 5 scale (ranging from unimportant to important), to guide the subordinate whether the discount rate should be below, at, or above the industry average (*DRG*), and to provide free-form written guidance (*WG*). Thus,

---

[10] The extent to which and conditions under which partner preferences trickle down to subordinates in practice are important empirical questions. Our experiment provides one test of this. The subordinates were unaware of supervisor preferences (and all supervisors were aware of this) so that we cleanly test how and where supervisors inject directional goals into audit team judgments. As our theoretical development of H1 discusses, this is more likely to occur via intervention cognition influencing the supervisors' finalized judgments than via overtly biased intervention contents, which would threaten justifiability.

[11] Providing supervisors and subordinates with the same case information at this stage allowed them both to reach the intervention and subordinate input stages at approximately the same time. This helped maintain control over the administration of the experiment. It also helped ensure that supervisors' ratings of their technical knowledge advantage were not contaminated by instrument-driven differences.

[12] A strong manipulation is consistent with the tenet to "design, plan, and conduct research so that the experimental conditions are as different as possible" (Kerlinger and Lee 2000, 459). In practice, many variants of active intervention exist, with some being more active and others less active than our treatment condition. Extant theory would predict such variation to influence the magnitude but not the direction of the effect of active intervention on supervisors' judgment bias.

intervening supervisors provided guidance for the subordinate's discount rate with some supporting reasons for the guidance.

We decided to structure the intervention in this manner to help ensure that participants assigned to the intervention condition would engage in more active intervention cognition than would supervisors not sufficiently motivated to intervene. However, structuring the intervention in this way limits our ability to separate the effects of the particular way in which the instrument prompted intervention cognition from the intervention itself. For example, we cannot isolate the extent to which the effects we observe are driven by prompting supervisors for importance ratings of the details relevant to their judgment, for directional rate guidance (*DRG*), or for free-form written-guidance (*WG*). Instead, we can only observe whether the active intervention as we operationalized it (including all of these prompts) is sufficient to show an effect. We discuss additional limitations of this design choice in Section VII.

Higher (lower) values of *DRG* indicate higher (lower) directional rate guidance. WG_{high} equals the number of written points favoring a higher rate, while WG_{low} equals the number of written points favoring a lower rate. *WG* equals the "net" written guidance (i.e., $WG = WG_{high} - WG_{low}$).[13] We also created *Index*, a composite summarizing the contents of a supervisor's intervention (including both directional and written guidance). Specifically, *Index* is the first principal component score from a factor analysis of *DRG*, WG_{high}, and WG_{low}.

Subordinates in the intervention condition then received their supervisor's guidance, formed their preliminary discount rate judgment (*SubJgmt*), and assessed their own task-relevant technical knowledge. As in the no-intervention condition, supervisors then received the subordinates' judgments (*SubJgmt*) and self-assessed knowledge, and then finalized a judgment for the audit team (*FinalJgmt*).

Supervisors' Perceived Knowledge Advantage

We measured supervisors' perceived technical knowledge advantage over subordinates by having each supervisor rate whether they believe their own technical knowledge dominates, approximates, or is dominated by that of the subordinate. We define k as supervisors' perception of their own technical knowledge relative to subordinates'. For interpretive reasons discussed in the results (footnote 15), we also define K as the mean-centered transformation of k (i.e., $K = k - \bar{k}$). Larger values of k and K indicate more perceived technical knowledge relative to the subordinate.[14]

---

[13] WG_{high}, WG_{low}, and *WG* are the averages of scores assigned by two content-coders: an author who was blind to the remainder of the data when content-coding the supervisors' interventions, and a doctoral student who was not blind to the remainder of the supervisors' data, but was unaware of the hypotheses and the subordinates' data. The two content-coders' scores are highly positively correlated ($r = 0.89$, $p < 0.001$ for WG_{high}; $r = 0.88$, $p < 0.001$ for WG_{low}; and $r = 0.91$, $p < 0.001$ for *WG*). Cronbach's α indicate very high inter-rater reliability ($\alpha = 0.93$ for WG_{high}, $\alpha = 0.93$ for WG_{low}, and $\alpha = 0.94$ for *WG*).

[14] We measured k at this point in the task because our theory would predict supervisors' *ex ante* impressions of their technical knowledge advantage to influence their judgments. We made this design choice because, had we measured it in a post-test questionnaire, our tests of H2 would have been confounded by a reverse causality explanation that supervisor participants altered their post-judgment assessments of k *ex post* to justify the extent to which goals influenced their judgments (MacKinnon [2008] and Kenny [2008] point out that experiments should measure such variables first when they would be tainted by such reverse-causality). However, we acknowledge that measuring k as we did can modify behavior. We believe that supervisors likely consult their perceptions of the technical knowledge of subordinates relative to themselves when getting their input on a technical matter, and we made our measurement as innocuous as possible. Supervisors in both conditions assessed k after assessing their own technical knowledge and receiving their subordinates' self-assessed technical knowledge. We anticipated that supervisors' perceived knowledge advantage over subordinates (k) would be, on average, a self-serving re-interpretation of their own and their subordinates' self-assessments. Results suggest that this is the case: supervisors' perceived knowledge advantage is positively related to the difference between the supervisors' and subordinates' self-assessments ($p = 0.048$), but self-servingly re-interpreted. Specifically, on average, supervisors perceived a significant technical knowledge advantage over their subordinates ($p =$

Dependent Variable and Post-Test Questionnaire

Supervisors then finalized the discount rate for the audit team (*FinalJgmt*), our primary dependent variable of interest. Finally, both supervisors and subordinates placed all of the experimental materials up to that point in an envelope and then completed a post-test questionnaire.

IV. EXPERIMENT 1: RESULTS

Manipulation Checks

While supervisors in the intervention condition could provide guidance, those in the no-intervention condition had no means of communicating with their subordinates. These features of the experiment gave us direct control over ensuring a successful intervention manipulation. Additionally, the post-test questionnaire asked all supervisors and all subordinates to rate how much they perceived that the supervisor's intervention influenced the subordinate's judgment. The auditors rated this on a scale from 0 ("there was no intervention; the subordinate made an independent assessment") to 10 ("a high amount of intervention; the guidance strongly influenced the subordinate's assessment"). Mean perception ratings are higher in the intervention condition than in the no-intervention condition for subordinates (5.3 versus 3.4, $t = 2.22$, $p_{one\text{-}tailed} = 0.017$) and weakly higher for supervisors (5.2 versus 3.6, $t = 1.54$, $p_{one\text{-}tailed} = 0.067$).

Comprehension Checks

At the beginning of the materials, two questions measured whether participants understood that (1) a decrease in the discount rate increases the fair value of the residual interest, and (2) an increase in the fair value of the residual interest increases the client's gain. Of the 29 supervisors, 27 (93 percent) answered the first question correctly and 26 (90 percent) answered the second question correctly. Of the 29 subordinates, 29 (100 percent) answered the first question correctly and 24 (83 percent) answered the second question correctly. Overall, it appears that the auditors generally understand the relationships between the discount rate and the client's valuation and gain.

Replicating the Effects of Supervisors' Intervention Content on Subordinates' Judgments

We examined whether the directional contents of supervisors' intervention influenced their subordinates' judgments, consistent with prior studies (e.g., Peecher 1996; Wilks 2002). Comparing our study to prior research, two differences are noteworthy: First, only half of our subordinates received intervention content. This leaves us with a modest sample of $n = 14$ audit teams to observe whether the directional content of the interventions directionally influences subordinates' judgments, replicating prior research. Second, prior studies actually manipulated the intervention content from hypothetical supervisors (who favored either more conservative or more aggressive accounting estimates) to test for its directional influence on subordinates' judgments, using experimenter-provided language that was held constant within conditions (e.g., Peecher 1996; Wilks 2002). Our experiment let actual supervisors provide their own language to subordinates, allowing for a replication and robustness test from experimenter-provided (prior studies) to participant-provided intervention content (this study).

0.037), even though the difference between supervisors' and subordinates' self-assessments did not really indicate to supervisors such an advantage ($p = 0.21$). We originally included supervisors' self-serving perceptions of their own knowledge advantage (k) in the instrument (rather than rely on the computed difference between the self-assessments) because our theory would predict the former to be more important than the latter in influencing supervisor judgment. Interestingly, and theory-consistent, we find supervisors' perceived knowledge advantage to be significant in our tests of H2 (see Section IV), but not the computed difference in the self-assessments of knowledge, even though the two are positively correlated.

We find evidence at $p < 0.06$ that supervisors' combined directional rate guidance and written guidance directionally influenced subordinates, replicating prior research. First, recall that we defined *Index* as the first principal component of *DRG*, WG_{high}, and WG_{low}, and, as such, *Index* is a composite measure of each supervisor's intervention content. The factor analysis (Table 3, Panel A) yields *Index* as a weighted average of *DRG*, WG_{high}, and $-WG_{low}$. These signs of the factor loadings suggest that the first principal component (which summarizes 65 percent of the variance in *DRG*, WG_{high}, and WG_{low}) captures the supervisors' collective directional guidance. Second, as Table 3, Panel B shows, *SubJgmt* is weakly positively correlated with *Index,* as well as with *DRG* and *WG*. Specifically, subordinates' discount rates are lower (higher) when supervisors suggested lower (higher) rates (*DRG*) and as supervisors provided fewer (more) items of written guidance favoring higher rates.

Tests of Hypotheses

Hypothesis 1 predicts that a supervisor's intervention distorts his or her finalized judgment for the audit team, beyond those of supervisors who have the same goals but do not intervene, and after controlling for any effects on the anchor provided by the subordinate's judgment. In contrast, H2 predicts that a supervisor's perceived knowledge advantage tempers the directional effect of goals on the supervisor's finalized judgment.

We use the following linear regression model to test our hypotheses:

$$FinalJgmt = \beta_0 + \beta_1 I + \beta_2 G + \beta_3 K + \beta_4 (I \times G) + \beta_5 (I \times K) + \beta_6 (G \times K) + \beta_7 (I \times G \times K) + \beta_8 SubJgmt + \varepsilon. \quad (1)$$

TABLE 3

Experiment 1

Effects of Supervisor's Intervention on Subordinates' Judgments

Panel A: Factor Analysis of Supervisor's Intervention[a]

	Index (First Principal Component)
DRG factor coefficient	1.36
WG_{high} factor coefficient	1.12
WG_{low} factor coefficient	−0.21
% variance explained	65%
n	14

Panel B: Correlations with Subordinates' Judgments (*SubJgmt*)[b]

	Predicted Sign	Pearson's r	p	Spearman's r	p	Kendall's τ	p
DRG	+	0.34	0.118	0.42	0.066	0.34	0.060
WG_{high}	+	0.45	0.055	0.54	0.024	0.34	0.062
WG_{low}	−	0.06	0.840	0.04	0.879	0.01	0.952
WG	+	0.39	0.086	0.37	0.094	0.31	0.077
Index	+	0.45	0.054	0.44	0.058	0.32	0.059

[a] Variable definitions appear at Table 2.

[b] Except for those of WG_{low}, the correlation estimates are in the predicted direction and p-values are one-tailed.

FinalJgmt is the supervisor's finalized judgment of the most reasonable discount rate for the audit team (*lower* values indicate more aggressive, client-preferred judgments). $I = 1$ for audit teams in the intervention condition, and 0 otherwise. G is the supervisor's mean-centered perception of the partner's preferences (*higher* values indicating more aggressive, client-preferred preferences), while K is the supervisor's mean-centered perceived task-relevant technical knowledge relative to the subordinate (*higher* values indicating a larger perceived knowledge advantage over the subordinate). *SubJgmt* (the subordinate's preliminary judgment) is a control variable. A classic anchoring effect (Tversky and Kahneman 1974) would predict the coefficient for *SubJgmt* to be significantly positive. Including *SubJgmt* allows for a cleaner test of H1 by controlling for any concurrent effects impounded in the subordinate's initial judgment. We test the above model and a parsimonious model that drops the two unhypothesized and insignificant terms $I \times K$ and $I \times G \times K$ in Table 4.[15]

As Table 4, Panel A indicates, the coefficient for *SubJgmt* is significantly positive ($p_{\text{one-tailed}} \leq 0.039$), suggesting that subordinates' judgments (and any effects impounded therein) are indeed reflected in the supervisors' finalized judgments for the audit team.[16]

Hypothesis 1 predicts that the coefficient for $I \times G$ will be negative. In other words, G will influence *FinalJgmt* more for intervening supervisors than for non-intervening supervisors, after controlling for any effects impounded in *SubJgmt*. Larger values of G indicate that the supervisor perceives that the engagement partner prefers more aggressive, client-preferred valuations. Lower values of *FinalJgmt* indicate lower discount rates and, therefore, more aggressive valuations.

As predicted by H1, the coefficient for $I \times G$ is significantly negative ($p_{\text{one-tailed}} \leq 0.016$). This result supports H1. This negative interaction indicates that the same goals to reach a preferred conclusion influenced supervisors' finalized discount rates more among supervisors who intervene than among those who do not, controlling for effects concurrently impounded in subordinates' judgments. When $I = 1$, the coefficient for G on *FinalJgmt* is weakly negative in the full model ($p_{\text{one-tailed}} = 0.056$) and significantly negative in the parsimonious model ($p_{\text{one-tailed}} = 0.033$; Table 4, Panel B). When $I = 0$, the coefficient for G on *FinalJgmt* is statistically insignificant ($p > 0.10$; Table 4, Panel A).

Hypothesis 2 predicts that supervisors' perceived knowledge advantage K will counteract the effects of G on *FinalJgmt*. As shown in Table 4, Panel A, the coefficient for $G \times K$ is significantly positive ($p_{\text{one-tailed}} < 0.001$). This interaction indicates that supervisors selected a discount rate

---

[15] We use the mean-centered variables G and K rather than g and k for the following reason: While our independent construct of interest are g and k, mean-centering these two continuous variables is necessary to properly interpret the t-tests in the above regression model as tests of H1 and H2 (Echambadi and Hess 2007). Mean-centering continuous variables is the equivalent of switching from a test of simple effects to a test of main effects in classical ANOVA (Echambadi and Hess 2007; Irwin and McClelland 2001). For example, if we had used g and k in the above model, then our test of H1 in Table 4 would show that intervention amplifies directional goals *when* $k = 0$ (Irwin and McClelland 2001). Using the mean-centered variables G and K appropriately switches the test of H1 in Table 4 to a test of at $I \times G$ *across the average of all levels of* k rather than at a single value that is not representative of the distribution of k (Irwin and McClelland 2001; Echambadi and Hess 2007). Correspondingly, H1 predicts intervention and goals to interact not just at one particular value of k, but across all levels of k. Mean-centering accomplishes this without increasing the risk of Type I errors (Echambadi and Hess 2007). We continue to use simple-effects coding of the categorical variable intervention (I) for exposition (e.g., Table 4, Panel B). When we re-run Table 4 using main effects coding for I (i.e., $I = 1$ or -1, as in classical ANOVA tests), those results are statistically similar to Table 4 and similarly support our hypotheses. Following Echambadi and Hess (2007), we test our regression models for multicollinearity using the uncentered variables that represent the constructs of interest (g and k) and multiple diagnostic tests. These tests indicate that we do not have multicollinearity problems (e.g., all VIFs < 10).

[16] Prior research indicates that supervisors are influenced by their subordinates' inputs (e.g., Ricchiute 1999). This effect has primarily been shown under conditions in which supervisors were relying on subordinates for information. We thank an anonymous reviewer for pointing out that our study extends this prior work by testing whether supervisors *still* rely on their subordinates' judgments as an input even when they are aware that they have approximately the same set of information as subordinates.

TABLE 4

Experiment 1

General Linear Regressions of Supervisors' Finalized Judgments

$$FinalJgmt = \beta_0 + \beta_1 I = \beta_2 G = \beta_3 K = \beta_4(I \times G) + \beta_5(I \times K) + \beta_6(G \times K) + \beta_7(I \times G \times K) + \beta_8 SubJgmt + \varepsilon$$ [a]

Panel A: Regression Analysis

Variable	Predicted Sign[b]	Coefficient (t- statistic)	Coefficient (t- statistic)
Intercept		8.63	8.76
		(10.10) ***	(10.41)***
I		0.16	0.19
		(0.64)	(0.74)
G		0.13	0.14
		(1.61)	(1.71)
K		0.43	0.38
		(4.05) ***	(4.04)***
$I \times G$	−(H1)	−0.26	−0.28
		(−2.32) **	(−2.66)***
$I \times K$		−0.26	—
		(−1.17)	—
$G \times K$	+ (H2)	0.29	0.29
		(6.18) ***	(6.58)***
$I \times G \times K$		−0.14	—
		(−0.89)	—
SubJgmt	+	0.16	0.15
		(1.98) **	(1.85)**
R^2		0.884	0.868
Adjusted R^2		0.836	0.830
n		28[c]	28[c]

Panel B: Supplementary Analysis

Variable	Predicted Sign[b]	Coefficient (t- statistic)	Coefficient (t- statistic)
G when $I = 1$	—	−0.13	−0.15
(i.e., $\beta_2 + \beta_4$)		(−1.67)*	(−1.94)**

*, **, *** Indicate $p < 0.10$, $p < 0.05$, and $p < 0.01$, respectively.

[a] Variable definitions appear at Table 2.

[b] Directional predictions are tested with one-tailed t-tests.

[c] One sample unit is excluded due to missing data.

more inconsistent (consistent) with their directional goals as their perceived knowledge advantage increased (decreased). This result supports H2.[17] Finally, Table 4 shows that the models explain a

[17] In supplementary analyses of H2, we mean-split *K* into a dichotomous variable to gain a better understanding of the nature of this $G \times K$ interaction. Results of these tests indicate that H2 is driven primarily by higher-knowledge

high percentage of the variance in supervisors' finalized judgments (adjusted R^2 approximately 83 percent).[18]

Supplementary Analyses and Discussion

As noted in the theoretical development of H1, people engaged in motivated reasoning generally do not consciously bias their judgments. Rather, they bias their judgments in subtle, unconscious ways that maintain an appearance of objectivity (Kunda 1990; Kadous et al. 2003; Piercey 2009). Because of this, H1 does *not* predict intervening supervisors to bias their intervention content in order for active intervention—via intervention cognition—to influence the discount rates they choose. Moreover, it is not necessary that intervention content be biased by directional goals in order for these goals to distort intervention cognition in a goal-consistent direction. Regardless of how intervention *content* concurrently affects subordinates, intervention *cognition* is predicted to amplify the effects of an intervening supervisor's own directional goals on the supervisor's final judgments (H1).

We find evidence of this causal explanation in supplementary analyses. First, the post-test questionnaire provides evidence that the intervention effects observed in our tests of H1 are unconscious. It asked all supervisors to rate the extent to which they believed that the engagement partner's views influenced their judgments on a scale from 0 to 10. Supervisors indicated that the partner's preference had at least some influence on their judgments ($t = 7.39$, $p < 0.001$). However, this did not vary by intervention condition (3.09 versus 2.93, $t = 0.18$, $p = 0.86$). This suggests that supervisors were unaware that intervention increased the effect of directional goals on their final judgments.[19] Second, while subordinates were indeed influenced by the direction of the guidance they received (replicating prior studies, Table 3), supervisors did not systematically tilt their intervention content in a goal-consistent direction. That is, we do not detect significant correlations between *G* and any of *DRG*, *WG*, *Index*, or *SubJgmt* among our 14 intervention audit teams (all $p \geq 0.35$).[20]

Thus, it is not simply the case that supervisors bias their judgments by steering their subordinates in the supervisors' goal-consistent directions and then simply anchoring on their subordinates' judgments, as suggested by prior research. The additional bias in the finalized judgments of intervening supervisors cannot simply be attributed to effects their intervention content had on

---

supervisors tempering their directional goals. Specifically, while all supervisors provide more goal-consistent discount rates as they coach subordinates ($t = -3.08$ $p = 0.006$), high-knowledge supervisors provide less goal-consistent discount rates ($t = 4.34$, $p < 0.001$). Low knowledge supervisors, on the other hand, do not ($t = 0.83$, $p = 0.43$).

[18] We interpret our high adjusted R^2 as indicative that the model variables explain systematic versus residual variance *within the experimental task* well, controlled extraneous sources of variance well, or both (Kerlinger and Lee 2000). We do not interpret this as indicating that the independent variables explain similarly high amounts of variance in the real world, for several reasons. Other sources of variance in the real world (e.g., interpersonal social cues that would accompany intervention) were controlled within the experimental setting but are not in the real world. Additionally, the experiment measures variables of interest to the experimental task with response scales designed to control variance. Overall, the model's adjusted R^2 suggests successful experimental design and good predictive validity within the experimental setting (Neter et al. 1996; Maxwell and Delaney 1990; Kerlinger and Lee 2000).

[19] Consistent with this, supervisors rated their perceptions of the accuracy of, their confidence in, and the justifiability of their judgments. We find no significant differences in these ratings between the intervention and no-intervention groups, further supporting the notion that H1 is an unconscious effect. In contrast, H2 predicts that a larger perceived knowledge advantage will lead supervisors to reduce their directional goals, which suggests that knowledge effects in H2 can be conscious. Consistent with this, supervisors' accuracy, confidence, and justifiability ratings *do* increase with perceived technical knowledge advantage ($r = 0.29$, $p = 0.064$; $r = 0.48$, $p = 0.004$; $r = 0.28$, $p = 0.072$, respectively).

[20] Prior research uses hypothetical supervisors and implicitly assumes that the supervisor's intervention content (i.e., *DRG, WG*) and the supervisor's directional goals (i.e., *G*) represent the same construct (e.g., Peecher 1996; Brown et al. 1999; Wilks 2002; Piercey 2009). We improve upon prior literature by using actual audit teams, an experimental design that allows us to disentangle these two constructs, and a causal model that acknowledges them as distinct constructs that can even be uncorrelated as supervisors with directional goals attempt to maintain an appearance of objectivity.

their subordinates. It also cannot simply be attributed to intervening supervisors adjusting for a subordinate's failure to be influenced by intervention in the supervisor's goal-consistent directions. Because both groups of supervisors received subordinate input uninfluenced by *G*, the effects of H1 are not due to differences in the subordinate's input. Non-intervening supervisors' subordinates were also not influenced by intervention, and non-intervening supervisors with the same goals could have made the same adjustment to the finalized discount rate, but did not (Table 4). Rather, these supplemental findings underscore the significance of H1; intervention cognition appears to cause supervisors to bias their judgments more than non-intervening supervisors with the same goals to reach a preferred conclusion, after controlling for any effects of intervention on the initial anchor provided by subordinates. These supplemental findings, together with our tests of H1, also highlight that the subordinate's judgment (the primary focus of prior research) is only an input into the audit team judgment. We extend prior literature with new evidence that intervention effects on the supervisor introduce a distinct bias into audit team judgments, separate from subordinate effects.

V. EXPERIMENT 2: METHOD

Motivation

Whether these effects are conscious or unconscious is a crucial determinant of the likelihood that auditors can successfully attempt to improve their own or other auditors' judgments via conscious review or quality-control procedures (Lerner and Tetlock 1999; Wegener et al. 2001; Moore and Loewenstein 2004). If auditors are unaware of the biasing effects of a supervisor's intervention on the supervisor, then training or policy changes would likely help pre-empt or correct the biasing effects of a supervisor's intervention. Training interventions could include making auditors aware of this effect and encouraging them to consider their own directional goals or striving to obtain relatively more independent work from their subordinates. Policy interventions could include increased emphasis on reviews by individuals who are more removed from the audit, such as the concurring or cold partner reviews that the PCAOB currently deems insufficient in practice (PCAOB 2008). Thus, the findings from Experiment 1 not only inform prior auditing research and theory, but they could also have important policy implications for audit firms and audit regulators. The less aware auditors are of these effects shown in Experiment 1, the more informative and adverse these implications can be. This warrants an experiment designed more specifically to test auditors' awareness of these effects (cf. Wilks 2002, Experiment 2).

Participants

One hundred five auditors from a major accounting firm volunteered as participants. Unlike Experiment 1, our sample unit for this experiment is an individual auditor. On average, participants had approximately three years of audit experience and 1.3 years as supervisors. Six participants did not complete the experimental instrument, giving us a sample of 99 auditors.

Design

We use a within-subjects design: a $3 \times 2 \times 2$ balanced, incomplete-block design (see Winer 1971, 622). The manipulated factors are *G* (goals to reach a more conservative, neutral, or aggressive conclusion), *I* (intervention or no intervention), and *K* (knowledge advantage or no knowledge advantage). These factors described the supervisor of a hypothetical auditing team. Each participant received a block of six of the 12 possible cases, and, for each, judged the amount of bias that would be present in the audit team judgments.

We chose a within-subjects design because the objective of the experiment is to measure auditors' intentionally expressed beliefs about how *G*, *I*, and *K* influence the amount of bias in

audit team judgments (Libby et al. 2002).[21] The incomplete block-design kept the task short.

The variable *G* manipulates the supervisor's goals to reach a preferred audit conclusion (in the same asset securitization context). The neutral conditions described the supervisor as preferring the best possible or highest quality accounting policies, without prejudice to earnings management objectives. The aggressive conditions described the supervisor as encouraging clients to use aggressive accounting practices. The conservative conditions describe the supervisor as encouraging clients to use conservative accounting practices.

The variable *I* manipulates whether the supervisor actively intervened in the subordinate's judgment. In the intervention conditions, the supervisor explained his expectations for the estimate, and guided the subordinate as s/he (the subordinate) formed a preliminary judgment. In the no-intervention conditions, the subordinate developed his/her judgment without any guidance from the supervisor. In all conditions, the subordinate then documented his/her judgment and forwarded this documentation to the supervisor. The supervisor then reviewed the subordinate's memo and finalized the audit team's estimate of fair value by making any adjustments deemed necessary to the subordinate's judgment. Note that this narrative describing the audit teams parallels what took place among actual audit teams in Experiment 1.

K manipulated the supervisor's technical knowledge relative to the subordinate. The knowledge-advantage (no-knowledge advantage) conditions described the supervisor as possessing tremendous (moderate) technical accounting expertise and as quite (somewhat) knowledgeable about accounting for asset securitizations. In comparison, all conditions described the subordinate as possessing moderate technical accounting expertise and as being somewhat knowledgeable about asset securitizations.

Task

Before commencing, participants read a description of a fictional client's asset securitization transaction (the same context as in Experiment 1) and answered a fill-in-the-blank question designed to test their understanding of how estimates of the fair value of residual interest influence reported income.

Dependent Variable and Post-Test Questionnaire

For each case, participants rated the extent to which bias causes the audit team's estimate of fair value to deviate from the best possible estimate. These ratings were on a seven-point scale from −3 ("Audit team estimate likely is extremely low") to 0 ("Audit team estimate undoubtedly is the best possible estimate") to +3 ("Audit team estimate likely is extremely high"). This rating is our primary dependent variable of interest for Experiment 2, *Perceived Bias*. Upon completing the cases, participants placed their answers in an envelope and completed a post-test questionnaire, which included manipulation checks and demographic questions.

VI. EXPERIMENT 2: RESULTS

Manipulation Checks

In post-test questions, we replicated language used in different levels of our manipulations and asked participants to evaluate the language along five-point scales. For *G*, the scale ranged from 1 ("The manager's preferences are highly biased toward conservative accounting estimates") to 5 ("The manager's preferences are highly biased toward aggressive accounting estimates").

[21] As Libby et al. (2002) note, within-subjects designs are particularly desirable for experiments that investigate individuals' current beliefs or thinking about how manipulated factors directionally affect phenomena of interest. Within-subjects designs also create a high-power environment for testing the presence or absence of intentionally expressed, current beliefs (Neter et al. 1996).

Mean ratings are 1.36 for the conservative level, 2.76 for the neutral level, and 4.78 for the aggressive level, and each change is significant (all $p < 0.001$). For *I,* the scale ranged from 1 ("There was no manager intervention") to 5 ("There was an extreme amount of intervention"). Mean ratings are 2.49 for the no-intervention level and 4.13 for the intervention level, a significant increase ($p < 0.001$). For *K,* the scale ranged from 1 ("The senior's technical accounting knowledge is much greater than the manager's") to 5 ("The manager's technical knowledge is much greater than the senior's"). The mean ratings are 3.54 for the no-knowledge-advantage level and 4.71 for the knowledge advantage level, a significant increase ($p < 0.001$).

Comprehension Checks

As mentioned previously, participants responded to pre-test, fill-in-the-blank questions that tested their understandings of bias, noise, and whether the effects of increases in the estimated fair value of retained interest affect income. Of the 99 participants, 95 (96 percent) answered the bias question correctly, while, for the accounting question, 88 (89 percent) answered correctly, seven answered incorrectly, three answered incoherently, and one did not respond. Similar to Experiment 1, participants generally appeared to understand the earnings implications of the estimated fair value of residual interest.

Perceived Bias in Supervisors' Finalized Judgments

An analysis of the variance of *Perceived Bias* (untabulated) yields only significant main effects of *G* ($F = 93.3$, $p < 0.001$) and *K* ($F = 8.26$, $p = 0.004$). With respect to *G,* the least-squares mean *Perceived Bias* ratings are 1.86 for the conservative conditions, 0.71 for neutral conditions, and 3.15 for the aggressive conditions. These bias ratings for the two directional goal conditions (conservative and aggressive), averaged together, are significantly higher than those of the neutral goal conditions ($t = 11.61$, $p_{\text{one-tailed}} < 0.001$). These findings suggest that auditors believe that directional goals to reach a preferred conclusion bias supervisors' finalized judgments. With respect to *K,* the least-squares mean bias ratings are 2.12 for the no-knowledge-advantage conditions and 1.70 for the knowledge-advantage conditions, a significant decrease ($t = 2.87$, $p_{\text{one-tailed}} = 0.002$). Thus, auditors appear to believe that a supervisor's knowledge advantage reduces bias in supervisors' finalized judgments.[22] With respect to *I,* we find no evidence that auditors believe that intervention increases bias in supervisors' finalized judgments. The least-squares mean bias ratings are 1.84 in the no-intervention conditions and 1.98 in the intervention condition ($F < 1$). Thus, auditors do not appear to believe that intervention influences bias.

VII. CONCLUSION

Results from the two experiments collectively support several conclusions with implications for both theory and practice. First, consistent with a concern raised in prior research that supervisor interventions distort subordinates' judgments, we find evidence that subordinates' judgments are reflected in the finalized judgment for the audit team. Thus, supervisors can influence and then anchor on subordinates' judgments. More relevant to our research question, intervention leads to distortions in the supervisor's finalized judgment for the audit team *beyond* any concurrent effects on the subordinate's judgment. Specifically, after taking subordinates' judgments into account, supervisors who intervene reach even more goal-consistent conclusions than do supervisors who

[22] Note that this main effect in auditors' *Perceived Bias* in supervisors' finalized judgment is consistent with the effect (*G* × *K*) on the actual levels of reviewer's finalized judgments, observed in Experiment 1. Because Experiment 1 uses finalized judgment levels as the dependent variable, we infer the tempering effect of *K* on bias from its tempering the bias-inducing effects of goals to reach a preferred directional conclusion *G* (i.e., *G* × *K*) on *FinalJgmt* (cf. Kunda 1990). In Experiment 2, we measure *Perceived Bias* directly, and the main effect of *K* analogously suggests that auditors believe *K* to have a tempering effect on bias.

have the same goals but do not intervene. This suggests that intervention creates additional cognitive effects in the mind of supervisors, consistent with theory underlying H1.

Formally or informally, supervisors have the option to intervene relatively early in subordinates' judgment processes. One view regards supervisors as reducing the bias in subordinates' judgments, and therefore more active intervention would lead to more objective judgments. While prior research has focused on the effect of intervention on subordinates' judgments, we add to this body of research by considering a setting in which coaching or actively intervening in a subordinate's judgment generates bias in supervisors themselves as they finalize audit team judgments.

Our findings from Experiment 1 suggest that supervisor intervention increases motivated reasoning bias in audit team judgments through its effects on the intervening supervisor. Experiment 2 provides evidence that auditors likely do not understand these effects. A lack of understanding of these effects precludes a realistic likelihood that supervisors could consciously attempt to adjust for them (Wegener et al. 2001), short of training or policy changes. To counter this risk, it would be prudent for firms to train auditors about how intervention affects bias in audit team judgments. Supervisors could try to become conscious of these intervention effects, or more simply avoid unnecessary early intervention in their subordinates' judgments, allowing their input to be more independent. This advice is likely especially prudent when stronger sources of client pressure that could affect the supervisors' independence are present. Additionally, audit firms could rely more on reviews done by supervising auditors who are more removed from the decision making of the subordinates they review. At the partner level, these "cold" concurring partner reviews have this purpose. The PCAOB (2008) currently identifies such reviews as probably the most frequently cited audit deficiency in current practice, and is considering a proposed standard that would substantially increase importance on them as an audit quality control.

Our experimental findings also suggest that, as supervisors perceive that their technical knowledge dominates their subordinates more, they temper the influence of their directional goals to reach a preferred conclusion on audit team judgments. Moreover, auditors appear to understand that perceived technical knowledge advantages temper bias in supervisors' judgments; this effect and auditors' understanding thereof enable supervisors to better use such knowledge advantages as a cue to temper their directional goals, better fulfilling the quality control role of audit review (cf. Rich 2004; Kunda 1990).

Kennedy and Peecher (1997) document that supervisors tend to be overconfident in their subordinates' technical knowledge, and that this overconfidence increases as the knowledge gap between them increases (cf. Messier et al. 2008). Linking this observation to our theory and findings, those supervisors who tend to be more aware of the actual knowledge gap probably are more likely to act as a reasonableness check in audit team judgments as the actual technical knowledge gap increases. However, this is less likely to occur for the modal supervisor who overestimates subordinates' technical knowledge, and does so increasingly as the actual gap increases. Thus, less overconfidence in assessing the technical knowledge of one's subordinates is likely to be an important criterion for promotion of reviewing auditors. Since Experiment 2 suggests that auditors are *aware* of this effect, firm policies that identify (or train) supervisors with more accurate confidence in their subordinates' judgments are likely to make supervisors more sensitive to reasonableness constraints in auditing team judgments.

Like all empirical-experimental investigations, ours has its limitations. First, we do not examine all factors that affect the amount of bias in audit team judgments. Rather, we limit our discussion to speaking to the moderating consequences of supervisors' active intervention and perceived knowledge advantages on the effects that directional goals have on supervisors' judgments. Additional research should examine other factors that impose boundary conditions on these effects. Second, in Experiment 1, we used measurement rather than manipulation to achieve variation in supervisors' directional goals. We did so to conserve on the number of two-member,

industry-specialist auditor teams needed to conduct our study. Additionally, theory would not predict intervention or supervisor knowledge advantage to interact differently with supervisor goals depending on whether they are manipulated versus measured. Third, we could have achieved additional realism in the task by adding more discount rate benchmarks. However, our theory would not predict this to alter the conclusions we reach so we forwent such realism (cf. Cook and Campbell 1979; Peecher and Solomon 2001; Kachelmeier and King 2002). Fourth, while we do not find evidence in Experiment 2 that auditors believe that intervention affects bias in audit team judgments, we cannot rule out the possibility that such beliefs exist but were unobservable due to inherent limitations in experimental design, statistical power, or measurement error.

Finally, the way in which we operationalized active intervention imposes limitations on the inferences we can draw. Specifically, Experiment 1 used structured prompts to elicit both intervention content and cognition from supervisors in the intervention condition. We could have allowed supervisors to guide subordinates without the structured cognition prompts, leaving the extent of intervention cognition up to them. Future research could measure key attributes of free-form intervention cognition by supervisors as well as investigate the attributes that are more or less apt to lead to intervention-induced biases in supervisors' judgments. Deeper goal-consistent intervention cognition during free-form intervention, for example, would be more likely than shallow goal-consistent processing to lead to such bias.

Despite these limitations, we find evidence with implications for audit supervisors about the determinants of bias in audit team judgments. We show how active intervention and supervisors' perceived knowledge advantages over subordinates have countervailing effects in that the former creates bias while the latter mitigates bias in audit team judgments. A worrisome rub is that auditors appear to understand only the latter, bias-mitigating effect.

REFERENCES

Abdolmohammadi, M. J. 1999. A comprehensive taxonomy of audit task structure, professional rank, and decision aids for behavioral research. *Behavioral Research in Accounting* 11 (1): 51–92.

Agoglia, C. P., J. F. Brazel, R. C. Hatfield, and S. B. Jackson. 2008. The effect of risk of misstatement and workload pressure on the choice of workpaper review format. Working paper, Drexel University.

Bamber, E. M., L. Bamber, and J. Bylinski. 1988. A descriptive study of audit managers' working paper review. *Auditing: A Journal of Practice & Theory* 7 (Spring): 137–149.

Bazerman, M. H., D. A. Moore, P. E. Tetlock, and L. Tanlu. 2006. Reports of solving the conflicts of interest in auditing are highly exaggerated. *Academy of Management Review* 31 (1): 43–49.

Boiney, L. G., J. Kennedy, and P. Nye. 1997. Instrumental bias in motivated reasoning: More when more is needed. *Organizational Behavior and Human Decision Processes* 72 (October): 1–24.

Brazel, J. F., C. P. Agoglia, and R. C. Hatfield. 2004. Electronic versus face-to-face review: The effects of alternative forms of review on auditors' performance. *The Accounting Review* 79 (4): 949–966.

Brown, C. E., M. E. Peecher, and I. Solomon. 1999. Auditor's hypothesis testing in diagnostic inference tasks. *Journal of Accounting Research* 37 (Spring): 1–26.

Chartrand, T. L., and J. A. Bargh. 1996. Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology* 71 (3): 464–478.

Cook, T. D., and D. T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston, MA: Houghton-Mifflin.

Cuccia, A. D., K. Hackenbrack, and M. W. Nelson. 1995. The ability of professional standards to mitigate aggressive reporting. *The Accounting Review* 70 (2): 227–248.

Dechow, P. M., and C. Shakespeare. 2009. Do managers time securitization transactions to obtain accounting benefits? *The Accounting Review* 84 (1): 99–132.

Echambadi, R., and J. D. Hess. 2007. Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science* 26 (3): 438–445.

Fargher, N., D. Freizer, and K. Trotman. 2005. A field-based analysis of work paper review. *Auditing: A Journal of Practice & Theory* 24 (2): 85–110.

Gibbins, M., and K. Trotman. 2002. Audit review: Managers' interpersonal expectations and conduct of the review. *Contemporary Accounting Research* 19 (Fall): 411–444.

Hales, J. 2007. Directional preferences, information processing, and investors' forecasts of earnings. *Journal of Accounting Research* 45 (3): 607–628.

Harris, R. 2008. New Century double trouble: Bad KPMG, no CFO. *CFO.com* (March 28).

Irwin, J. R., and G. H. McClelland. 2001. Misleading heuristics and moderated multiple regression models. *Journal of Marketing Research* 38 (1): 100–109.

Kachelmeier, S., and R. King. 2002. Using laboratory experiments to evaluate accounting policy issues. *Accounting Horizons* 16 (3): 219–232.

Kadous, K., and A. M. Magro. 2001. The effects of exposure to practice risk on tax professionals' judgments and recommendations. *Contemporary Accounting Research* 18 (3): 451–475.

———, ———, and B. C. Spilker. 2008. Do effects of client preference on accounting professionals' information search and subsequent judgments persist with high practice risk? *The Accounting Review* 83 (1): 133–156.

———, J. Kennedy, and M. E. Peecher. 2003. The effect of quality assessment and directional goal commitment on auditors' acceptance of client-preferred accounting methods. *The Accounting Review* (July): 725–758.

Kennedy, J., and M. E. Peecher. 1997. Judging auditors' technical knowledge. *Journal of Accounting Research* 35 (Autumn): 279–293.

Kenny, D. A. 2008. Mediation. *DavidAKenny.net* (September 12).

Kerlinger, F. N., and H. B. Lee. 2000. *Foundations of Behavioral Research.* 4th edition. New York, NY: Thomson Learning.

Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108 (November): 480–498.

Lerner, J. S., and P. E. Tetlock. 1999. Accounting for the effects of accountability. *Psychological Bulletin* 125 (2): 255–275.

Libby, R., R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27 (November): 775–811.

———, and K. Trotman. 1993. Audit review as a control for biased recall of evidence in decision making. *Accounting, Organizations and Society* 18 (6): 559–574.

MacKinnon, D. P. 2008. *Introduction to Statistical Mediation Analysis.* Mahwah, NJ: Erlbaum.

Maxwell, S., and H. Delaney. 1990. *Designing Experiments and Analyzing Data.* Belmont, CA: Wadsworth.

Messier, W. F., V. Owhoso, and C. Rakovski. 2008. Can audit partners predict subordinates' ability to detect errors? *Journal of Accounting Research* 46 (5): 1241–1264.

Moore, D. A., and G. Loewenstein. 2004. Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research* 17 (2): 189–202.

———, P. E. Tetlock, L. Tanlu, and M. H. Bazerman. 2006. Conflicts of interest and the case of auditor independence: Moral seduction and strategic issue cycling. *Academy of Management Review* 31 (1): 10–29.

Nelson, M. W. 2006. Ameliorating conflicts of interest in auditing: Effects of recent reforms on auditors and their clients. *Academy of Management Review* 31 (1): 30–42.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. *Applied Statistical Linear Models.* Boston, MA: MCB McGraw-Hill.

Ng, T. B.-P., and P. G. Shankar. 2010. Effects of technical department's advice, quality assessment standards, and client justifications on auditors' propensity to accept client-preferred accounting methods. *The Accounting Review* 85 (5): 1743–1761.

Payne, E. A., R. J. Ramsey, and E. M. Bamber. 2010. The effect of alternative types of review on auditors' procedures and performance. *Auditing: A Journal of Practice & Theory* 29 (May): 207–220.

Peecher, M. E. 1996. The influence of auditors' justification processes on their decisions: A cognitive model and experimental evidence. *Journal of Accounting Research* 34 (1): 125–140.

———, and I. Solomon. 2001. Theory and experimentation in studies of audit judgment and decisions: Avoiding common research traps. *International Journal of Auditing* 5 (November): 193–203.

Petty, R. E., and D. T. Wegener. 1998. Attitude change: Multiple roles for persuasion variables. In *The Handbook of Society Psychology*, edited by Gilbert, D. T., S. T. Fiske, and G. Lindzey. 4th edition, 323–390. New York, NY: McGraw-Hill.

Phillips, F. 1999. Auditor attention to and judgments of aggressive financial reporting. *Journal of Accounting Research* 37 (Spring): 167–190.

Piercey, M. D. 2009. Motivated reasoning and verbal vs. numerical probability assessment: Evidence from an accounting context. *Organizational Behavior and Human Decision Processes* 108 (2): 330–341.

Public Company Accounting Oversight Board (PCAOB). 2008. *Proposed Auditing Standard on Engagement Quality Review*. February 26. Washington, D.C.: PCAOB.

Pyszczynski, T., and J. Greenberg. 1987. Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. In *Advances in Experimental Social Psychology*, edited by Berkowitz, L., Vol. 20. New York, NY: Academic Press.

Ricchiute, D. N. 1999. The effect of audit seniors' decisions on working paper documentation and on partners' decisions. *Accounting, Organizations and Society* 24 (2): 155–171.

———. 2006. *Auditing*. 8th edition. Cincinnati, OH: South-Western Publishing.

Rich, J. S. 2004. Reviewers' responses to expectations about the client and the preparer. *The Accounting Review* 79 (April): 497–517.

———, I. Solomon, and K. Trotman. 1997a. The audit review process: A characterization from the persuasion perspective. *Accounting, Organizations and Society* 22 (July): 481–505.

———, ———, and ———. 1997b. Multi-auditor judgment/decision making research: A decade later. *Journal of Accounting Literature* 16: 86–126.

Salterio, S., and L. Koonce. 1997. The persuasiveness of audit evidence: The case of accounting policy decisions. *Accounting, Organizations and Society* 22 (August): 573–587.

Securities and Exchange Commission (SEC). 2010. Public Company Accounting Oversight Board: Order Approving Proposed Rules on Auditing Standard No. 7, Engagement Quality Review, and Conforming Amendment. *Release No. 34–61363* (January 15). Washington, D.C.: Government Printing Office.

Shankar, P. G., and H.-T. Tan. 2006. Determinants of audit preparer's workpaper justifications. *The Accounting Review* 81 (March): 473–495.

Sherman, J. W., B. Gawronski, K. Gonsalkorale, K. Hugenberg, T. J. Allen, and C. J. Groom. 2008. The self-regulation of automatic associations and behavioral impulses. *Psychological Review* 115 (2): 314–335.

Tan, S.-K., and H.-T. Tan. 2008. Effects of exposure to subsequently invalidated evidence on judgments of audit workpaper preparers and reviewers. *Contemporary Accounting Research* 25 (3): 921–946.

Turner, C. 2001. Accountability demands and the auditor's evidence search strategy: The influence of reviewer preferences and the nature of the response. *Journal of Accounting Research* 39 (3): 683–706.

Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185: 1124–1130.

Vera-Muñoz, S. C., J. L. Ho, and C. W. Chow. 2006. Enhancing knowledge sharing in public accounting firms. *Accounting Horizons* 20 (2): 133–155.

Wegener, D. T., M. Dunn, and D. Tokusato. 2001. The flexible correction model: Phenomenology and the use of naive theories in avoiding or removing bias. In *Cognitive Social Psychology: The Princeton Symposium on the Legacy and Future of Social Cognition*, edited by G. B. Moskowitz, 277–290. Mahwah, NJ: Erlbaum.

Wilks, T. J. 2002. Predecisional distortion of evidence as a consequence of real-time audit review. *The Accounting Review* 77 (January): 51–72.

Winer, B. J. 1971. *Statistical Principles in Experimental Design*. New York, NY: McGraw-Hill.

Winograd, B., J. Gerson, and B. Berlin. 2000. Audit practices of PricewaterhouseCoopers. *Auditing: A Journal of Practice & Theory* 19 (2): 175–182.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1787–1807

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1787

Perceptions of Investment Risk Associated with Material Control Weakness Pervasiveness and Disclosure Detail

Jacob M. Rose
University of New Hampshire

Carolyn Strand Norman
Virginia Commonwealth University

Anna M. Rose
University of New Hampshire

ABSTRACT: This research examines whether investors adjust their assessments of investment risk in response to material control weakness disclosures, the pervasiveness of material control weaknesses, and the detail of explanation provided regarding the pervasiveness of material control weaknesses. Findings from a laboratory experiment with 97 nonprofessional investors, a second experiment with 53 nonprofessional investors, and surveys of 47 investors and 28 *Fortune* 500 directors confirm prior archival findings that investors adjust their investment risk assessments in response to material weakness disclosures. More importantly, we find evidence of an interactive effect of material control weakness pervasiveness and disclosure detail that is counter to the expected benefits of expanded disclosure desired by corporate directors. When material weakness disclosures include specific and detailed discussion of the pervasiveness of control weaknesses, investors increase assessments of investment risk for less pervasive weaknesses and decrease assessments of risk for more pervasive weaknesses. Results indicate that these findings are driven by different levels of investor trust in management.

Keywords: *internal control; control pervasiveness; control weakness; disclosure detail; investment risk; management credibility; trust.*

We thank Steve Kachelmeier and two anonymous reviewers for their thoughtful comments. We also thank James Bierstaker, Ed O'Donnell, the participants of the 2006 Accounting, Behavior and Organizations Research Conference, and participants of workshops at Southern Illinois University, Virginia Commonwealth University, and the University of South Florida for their valuable insights. The authors acknowledge a research grant from the School of Business, Virginia Commonwealth University.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: June 2008
Accepted: January 2010
Published Online: August 2010

I. INTRODUCTION

Section 404 of the Sarbanes-Oxley Act of 2002 (SOX, U.S. House of Representatives 2002) and associated SEC regulations (SEC 2003) require management of public companies to establish and maintain adequate internal control over financial reporting and to obtain an attestation report by an independent auditor. In addition to promoting investor confidence, the requirements of Section 404 of SOX are intended to provide investors with more accurate information regarding the integrity of financial reports. Further, internal control weakness disclosures are believed to be effective signals of risks to investors created by potentially unreliable financial reports, allowing investors to make better forward-looking assessments of investment risk (PCAOB 2007).

The objectives of this research are to examine whether nonprofessional investors adjust their assessments of investment risk in response to material control weakness (hereafter, control weakness) disclosures, the pervasiveness of control weaknesses, and the detail of explanations provided regarding the pervasiveness of control weaknesses. Existing archival research indicates that control weakness disclosures result in adverse stock price reactions (e.g., Beneish et al. 2008; Hammersley et al. 2008), providing evidence that investors perceive risks related to material control weaknesses.

The current research leverages the advantages of experimentation to extend archival findings and explore the underlying causes of the effects of pervasiveness and explanation detail on investors' judgments. We first examine the effects of the pervasiveness of control weaknesses on perceived investment risk. Pervasive weaknesses are those at the financial statement and entity level that might affect overall control objectives, while less pervasive, or isolated, control weaknesses are typically associated with specific accounts and disclosures (Kinney 1989; Messier and Austen 2000; PCAOB 2007). Isolated weaknesses affect one account or transaction and pervasive weaknesses have system-wide effects that are inherently less precise (i.e., more vague) (Doss 2004). Examples of pervasive weaknesses include items such as problems with the control environment and ineffective information technology controls that affect many accounts and processes. Both auditors and regulators have stated that investors may not react to control weakness disclosures if they cannot comprehend the effects of weaknesses on future performance (Orenstein 2004; Hammersley et al. 2008). Our study examines whether individual investors recognize the differential effects of internal control weaknesses that are isolated versus pervasive, and whether they adjust investment risk perceptions accordingly. We also examine the effects of the explanation detail provided in material weakness disclosures on risk assessments. Psychology-based theories of risk assessment suggest that increased explanation decreases perceived risk. However, attribution theory indicates that more detailed explanations can interact with weakness pervasiveness, and increased risk perceptions can result.

Consistent with archival research, our results indicate that nonprofessional investors' assessments of investment risk increase for firms with material control weakness disclosures, relative to firms without such disclosures. Investors also recognize the pervasiveness of the control weaknesses, but the effects of control weakness pervasiveness are moderated by the level of explanation provided regarding pervasiveness. We also find significant evidence for an effect that is counter to the expected benefits of detailed explanations currently desired by corporate directors. When material weakness disclosures in the Management Report on Internal Controls include specific and detailed discussion of the pervasiveness of the control weaknesses, investors *increase* their assessments of investment risk for less pervasive weaknesses and *decrease* their assessments of risk for more pervasive weaknesses. That is, providing a detailed and explicit explanation that a control weakness is isolated and not pervasive increases forward-looking risk assessments relative to not providing a detailed explanation. Conversely, providing a detailed and explicit explanation that a control weakness is very pervasive decreases forward-looking risk assessments rela-

tive to not providing a detailed explanation. Supplemental analyses provide an explanation for this interactive effect.

Risk perceptions are amplified when decision makers lose trust in the source of information (Kasperson et al. 1988; Pidgeon et al. 2002). Our nonprofessional investor participants appear to have believed that management was less trustworthy when detailed explanations were provided for isolated, non-pervasive control weaknesses. On the other hand, participants perceived that management was more trustworthy when detailed explanations were provided for pervasive control weaknesses. Apparently, nonprofessional investors believe that management is attempting to downplay the significance of the problem when detailed explanations are provided for isolated control weaknesses, but detailed explanations of pervasive control weaknesses are viewed as indicators that management is trustworthy.

These results indicate that providing explicit explanations of control weakness pervasiveness may have unintended consequences to the extent that investors use the explanations as cues for assessing their trust in management. Insofar as the intent of Section 404 was to facilitate the restoration of investor confidence in the financial markets and provide relevant information to assess investment risk, these results should be interesting to legislators, investors, managers, and directors contemplating the structure of control weakness disclosures, as well as to independent auditors. We find that investors use information about control weaknesses and their pervasiveness to adjust forward-looking risk assessments. Additionally, the level of explanation provided in control weakness disclosures significantly affects investors' interpretations of investment risk.

II. BACKGROUND AND DEVELOPMENT OF HYPOTHESES

Control Weaknesses and Investment Risk

The U.S. Congress required the Securities and Exchange Commission (SEC) to adopt rules to implement Section 404 of the Sarbanes-Oxley Act. Under these rules, management must disclose any material weakness and may not conclude that the company's internal control over financial reporting is effective if even one material weakness exists. Analyses of recent disclosures under SOX Section 404 suggest that a number of adverse effects are associated with internal control weaknesses. Hammersley et al. (2008) find that returns are significantly negative when material weaknesses are disclosed, and the adverse effects of control weaknesses on returns are more pronounced when management claims that the control system is effective but the independent auditor reports that the system is ineffective. The authors suggest that these results are consistent with investor concerns about the expense required to remediate weaknesses or the possibility that uncorrected control weaknesses remain. Similarly, Beneish et al. (2008) find a significant negative price response for disclosing firms that is exacerbated by conditions associated with higher inherent reporting risk, such as auditor turnover and high-risk industry membership. Finally, Ogneva et al. (2007) report marginally higher cost of equity for firms that disclosed a material weakness in internal controls and for firms that delayed their Section 404 disclosures.

Palmrose et al. (2004) state that material control weaknesses should adversely affect stock price because weaknesses lead to material misstatements, and these material misstatements in financial statements result in future stock price declines when the misstatements are corrected. The prior archival research supports this perspective, and indicates that investors recognize the investment risks created by internal control weaknesses (Ogneva et al. 2007; Beneish et al. 2008; Hammersley et al. 2008). The first hypothesis is designed to confirm this finding under controlled laboratory conditions with nonprofessional investors. Given that we examine nonprofessional investors, it is important to verify the market-based findings because market effects could be driven by institutional investors and other more sophisticated investors. We propose the following hypothesis to confirm the findings of prior archival research in a controlled laboratory environment.

H1: The forward-looking investment risk perceptions of nonprofessional investors are higher for firms with material control weakness disclosures relative to firms without material control weakness disclosures, *ceteris paribus*.

Control Weakness Pervasiveness

Pervasive control weaknesses are those weaknesses that influence multiple aspects of the financial reporting system, while less pervasive, or isolated, control weaknesses influence individual accounts or classes of transactions (Kinney 1989; Messier and Austen 2000; PCAOB 2007). Extremely pervasive control weaknesses, such as failures of the information systems that produce financial information, create ubiquitous weaknesses throughout the entire financial system. Neither existing archival nor experimental research has examined whether investors recognize the pervasiveness of control weaknesses or whether investors react differentially to varying levels of pervasiveness. Prior research finds that auditors react to the pervasiveness of risks (Messier and Austen 2000), but researchers have not addressed the ability of nonprofessional investors to use information about weakness pervasiveness in their assessments of risk.

Hammersley et al. (2008) find that the market reacts more to material weaknesses than to less severe control deficiencies. The authors suggest that material weaknesses are more likely to result in future restatements of misstated financial information than are less severe control deficiencies, and material weaknesses trigger greater concerns about cost of remediation and future control weaknesses. Similarly, we expect investors to react to the pervasiveness of material control weaknesses, because more pervasive control weaknesses have greater potential to cause misstatements than do less pervasive weaknesses, and the widespread nature of pervasive weaknesses increases the potential costs of remediation and likelihood of future discoveries of additional control problems. According to Weber (1988) and Weber and Bottom (1989, 1990), decision makers assess risk by analyzing the severity and likelihood of potential outcomes. Pervasive control weaknesses are more severe than are isolated weaknesses, and the likelihood of misstatements increases for pervasive, relative to isolated, control weaknesses. Accordingly, we expect that investors' perceptions of investment risk will increase for more pervasive control weaknesses, relative to less pervasive control weaknesses.

H2: The forward-looking investment risk perceptions of nonprofessional investors increase as the pervasiveness of material control weaknesses increases.

Explanation Detail

There are potentially competing effects of the level of explanation provided in control weakness disclosures on investors' risk assessments. Hammersley et al. (2008) find that vague disclosures of material weaknesses (i.e., those with insufficient explanation detail to understand the underlying cause, the accounts affected, and the severity of the weakness) result in more adverse market reactions than do detailed disclosures. This is consistent with psychological theories of risk perception, by which decision makers perceive more risk when a risky item is less observable and more is unknown about the item. Slovic (1987) suggests that risk assessments are derived from two components: "dread" and "unknown." Based upon Slovic's model of risk, when less is known about a material control weakness, perceptions of the risks associated with the control weakness will increase. Slovic's model of risk perception therefore suggests a main effect of explanation detail, with more explanation detail decreasing perceptions of investment risk by decreasing the amount of unknown information related to the weakness.

Corporate directors also subscribe to the theory that less detailed disclosures increase investors' risk perceptions. A 2005 survey of more than 1,100 corporate directors by PricewaterhouseCoopers (PWC 2005) indicates that corporate directors believe that pervasive control matters

should be clearly separated from those that are less pervasive to help investors gauge the seriousness of any weaknesses (Kerstetter and Schulken 2005). Directors appear to assume that nonprofessional investors are unable to differentiate between more and less pervasive control weaknesses, leading to an over-estimation of risk related to less pervasive weaknesses. We survey a sample of 28 corporate directors from *Fortune* 500 firms to confirm the assumption that one source of directors' concerns regarding the pervasiveness of control weaknesses involves investors' reactions to these disclosures. Our survey results indicate that directors believe that investors over-react (i.e., perceive more risk than is warranted) to disclosures of more isolated control weaknesses and that increasing the explanation detail provided in weakness disclosures could reduce unwarranted investor reactions. Directors believe that providing detailed explanations of control weakness pervasiveness should decrease risk perceptions for less pervasive control weaknesses.

Attribution theory, however, suggests an interactive effect of explanation detail and pervasiveness, rather than a main effect of explanation detail. In addition, Hammersley et al.'s (2008) definition of "vague" disclosures contains features of both pervasiveness and explanation detail. Thus, how these two aspects of control weaknesses and related disclosures affect investors remains an open question. According to attribution theory, decision makers evaluate messages (such as management's disclosures concerning control weaknesses) by comparing messages with the incentives of the messenger (Eagly and Chaiken 1993). Attribution theory indicates that recipients (e.g., investors) of disclosures from a messenger (e.g., management) develop expectations of the contents of disclosures, and recipients determine the validity of the disclosure and the credibility of the messenger based upon comparisons of expectations and actual disclosures (Eagly and Chaiken 1975). Mercer (2005, 725) defines management's credibility related to financial reporting as "investor's beliefs about management's trustworthiness and competence in financial disclosure." Thus, trust is a key component of credibility. If a message is consistent with the messenger's incentives, then recipients view the message as biased and lose trust in the messenger. Good news disclosures are inherently viewed more skeptically by recipients, while bad news disclosures are more readily accepted, because bad news disclosures are inconsistent with the messengers' incentives (Eagly and Chaiken 1993).

Managers have incentives to exaggerate good financial news and downplay bad news (Hutton et al. 2003). Earnings forecast research finds that investors believe that bad news forecasts are credible because management has few incentives to report bad news (Skinner 1994; Soffer et al. 2000). As a result, investors consistently view bad news disclosures as more credible than good news disclosures. However, managers are also more likely to create explanations for bad news that externalize the causes of bad news, and investors expect management to downplay bad news (Baginski et al. 2000).

Attribution research in psychology and accounting suggests that investors expect that bad news will be accompanied by attempts to marginalize the bad news, and explanations that downplay bad news are viewed as biased. Investors, however, do not expect management to provide detailed explanations that explain the serious nature of more pervasive weaknesses (i.e., they do not expect management to emphasize bad news), and investors are likely to attribute greater levels of trustworthiness to managers that reveal the serious nature of pervasive weaknesses. Research demonstrates connections between trust in management and investor risk-taking (e.g., Ryan and Buchholtz 2001; Das and Teng 2004; Schwarzkopf 2006). This research generally finds that investors must perceive an acceptable level of trustworthiness in management before they decide to take the risk of investing in that firm, and increased trust can result in decreased perceptions of investment risk. Thus, following attribution and trust theories, increased explanation detail will increase perceptions of risk for less pervasive weaknesses and will decrease perceptions of risk for more pervasive weaknesses.

Given the competing predictions of attribution theory and psychological theories of risk assessment, we test the following null hypothesis:

H3: Increasing the level of explanation for material control weaknesses will not affect perceived investment risk.

III. METHOD

Design and Participants

The experiment involved a 2 × 2 between-participants design with a control group included as a fifth experimental condition. The experiment manipulated the pervasiveness of material internal control weaknesses (low and high) and the detail of explanation of the control weaknesses (low and high). The control group case materials did not include any internal control weaknesses. This group was used to evaluate the effects of a material weakness disclosure on investors' investment risk perceptions relative to risk perceptions related to annual reports where no material weaknesses are identified (H1). Participation in the study was voluntary.

Participants were M.B.A. students from two large, public U.S. universities. Elliott et al. (2007) find that M.B.A. students are reasonable surrogates for nonprofessional investors, and M.B.A. students have been employed as surrogates for nonprofessional investors in multiple prior studies investigating investor behavior (e.g., Hirst et al. 1995; Maines and McDaniel 2000; Hodge et al. 2004). The participants had an average of 6.1 years of management experience; 54 percent were male. Participants were asked if they understood the Management Report on Internal Controls. On a Likert scale ranging from 1 (Not at All) to 9 (Completely), the average response was 6.08.

Task and Procedure

Participants were randomly assigned to treatment conditions. The experiment provided participants with general instructions, case materials, and a post-experiment questionnaire. The task was completed under controlled conditions and under the supervision of an experimenter. Participants read the case materials and responded to a series of questions about investment risk and trust. The case described a manufacturing company and included several financial ratios from the prior three years and quarterly stock price data from the prior eight quarters. The company description, ratio data, and stock price data were derived from an actual manufacturing company (the company name was changed for the experiment). The brief sets of ratio and stock price data were included to demonstrate that the company experienced a steady improvement in financial ratios and steady growth in earnings and stock price over the period described, creating a favorable impression of the firm and its associated investment risk prior to the control weakness disclosure.

After reading the case materials, participants received the Management Report on Internal Controls described in AS No. 5 (PCAOB 2007) and the auditor's opinion on the Management Report on Internal Controls. The management reports and the auditor reports followed the format described by AS No. 5. The reports were identical across treatment conditions, with the exception of the specific control weaknesses identified and the level of explanation of those weaknesses. The management report in the control condition did not contain any disclosures of material internal control weaknesses.

In the four treatment conditions with material weaknesses, management identified a control weakness in its report on internal controls and the auditor concurred with management's conclusions. As a result, the auditor's report was unqualified with respect to management's assertion that internal controls were not effective, and adverse with regard to the effectiveness of controls. Hence, there was no conflict between the findings of management and the auditor, and there was

no indication of a management failure to detect the control weaknesses. Therefore, any differences in participants' risk perceptions are due to the manipulations of the material control weaknesses and the level of explanation of those weaknesses.

Independent Variables

Both the low- and high-pervasiveness control weaknesses involved a lack of adequate password/access controls. The text in the opinion paragraph that described the material weakness is presented below for the low-pervasiveness and low explanation conditions:

Low pervasiveness (low explanation):

> The material weaknesses resulted from a failure to maintain adequate password and access controls over the portion of the Enterprise Resource Planning system that ensures proper recording of sales in the appropriate reporting period.

In the high explanation level, the following text was added to the end of the opinion paragraph:

Low pervasiveness (additional text for high explanation):

> Access and password controls over the system for recording sales prevent unauthorized use of the system and unauthorized changes to the system by employees, management, and outsiders. The system reviews sales and shipping transactions that occur near period-end. The access controls ensure that sales revenues are always recorded in the proper accounting period. The insufficient password and access controls in the system for recording sales can only affect the recording of revenue for specific sales transactions near period-end, and this deficiency cannot affect any other aspect of the financial reporting system.

The control weakness and explanation level for the high-pervasiveness condition are presented below:

High pervasiveness (low explanation):

> The material weaknesses resulted from a failure to maintain adequate password and access controls over the entire Enterprise Resource Planning system.

High pervasiveness (additional text for high explanation):

> The Enterprise Resource Planning system includes all financial and nonfinancial systems and databases within the organization. Password and access controls prevent unauthorized use of the system and unauthorized changes to the system by employees, management, and outsiders. Insufficient password and access controls over the enterprise resource planning system can influence the integrity of many aspects of both financial reporting systems and nonfinancial systems, and the weaknesses could have resulted in material misstatements throughout the financial statements.

Dependent Variables

After reading and analyzing the Management Report on Internal Controls and the auditor report expressing an opinion on internal controls, as well as an opinion on the management report, participants made forward-looking assessments of the risk of a moderate stock price decline (1–10 percent) within the next year (where 1 = very low, 5 = moderate, and 9 = very high). This

measure of investment risk captures an individual's risk of loss. We use the risk of a stock price decline as our measure of investment risk for two reasons. First, risk is typically viewed in both the accounting and psychology contexts as a risk of personal loss (e.g., Fischhoff 1995; Olsen 1997; Loewenstein et al. 2001). Stock price declines represent clear and measurable losses to investors. Second, we pre-tested multiple measures of investment risk with more than 100 M.B.A. students enrolled in an investments course, and the risk of a stock price decline was found to be the most representative of personal investment risk. We include a second dependent variable: "How will the stock price of Davis Manufacturing Company react to the control weakness disclosure??" (where −5 = Significant Decline, 1 = No Effect, and 5 = Significant Increase) to allow for the possibility that participants will believe that control weaknesses could favorably affect future performance.[1] Finally, we assess trust in management with the following question: "Do you trust the management of Davis Manufacturing Company to accurately represent the financial position of the firm?" (where 1 = Definitely No, 5 = Moderately, and 9 = Definitely Yes).

IV. RESULTS

Manipulation Checks and Preliminary Tests

The participants were asked about the presence/absence of a material control weakness. Four participants responded incorrectly (i.e., four participants who were in a treatment group with a control weakness responded that there was no control weakness, indicating that they did not attend to the experimental materials) and one response was an extreme outlier. These five participants are not included in any of the analyses. Regarding the manipulation check for the pervasiveness treatment, the difference in the mean response for the low-pervasiveness (5.18) conditions is significantly lower than the mean response for the high-pervasiveness (6.86) conditions ($p < 0.001$). Thus, the pervasiveness manipulation appears successful.

Analyses of the dependent variable (forward-looking assessments of the risk of moderate stock price decline) are shown in Table 1. Panel A of Table 1 provides means, standard deviations, and sample sizes for all treatment conditions (including the control group) and main effects for the risk of a moderate stock price decline. The pattern of means suggests a potential interactive effect of control weakness pervasiveness and explanation detail on investment risk perceptions, but no main effects are readily apparent. Next, we use a regression model to conduct preliminary tests, where the dependent variable is the assessment of risk of a moderate stock price decline. The independent variables are the control weakness pervasiveness, explanation detail, and an interaction term (pervasiveness × explanation); and the covariates include demographic factors (i.e., age, gender, management experience). None of the demographic factors are statistically significant; thus, they are not included in our reported analyses.

Tests of Hypotheses

Our first hypothesis (H1) proposes that forward-looking risk assessments of nonprofessional investors increase for firms with control weakness disclosures relative to firms without control weakness disclosures. We test this hypothesis with an ANOVA model that includes the four treatment conditions and the control group. The dependent variable reflects the participants' assessment of the risk of a moderate stock price decline, and the independent variable is the treatment group (five groups). The group variable is statistically significant in the model ($F = 1072.25$, $p < 0.001$). *Post hoc* mean comparisons using the Bonferroni procedure indicate that assessments of the risk of a moderate stock price decline are lower in the control group relative to each of the

[1] No participants indicated an improvement in performance resulting from the control weakness.

TABLE 1

Descriptive Statistics and ANOVA (Dependent Variable = Risk of Moderate Stock Price Decline)

Panel A: Descriptive Statistics

Mean (Standard Deviation) {Sample Size} Responses across Treatment Conditions

Pervasiveness[a]	Explanation Detail[b]		Main Effect: Pervasiveness
	Low	High	
Low	5.65	6.47	6.06
	(1.96)	(1.42)	(1.74)
	{17}	{17}	{34}
High	6.59	5.68	6.18
	(1.40)	(1.96)	(1.71)
	{22}	{22}	{44}
Main Effect: Explanation	6.18	6.05	6.10
	(1.72)	(1.75)	(1.73)
	{39}	{39}	{78}

Control Group: 3.74 (1.56) {19}

Panel B: ANOVA Model

Source	Sum of Squares	df	Mean Square	F	p-value
Pervasiveness (P)[a]	0.115	1	0.115	0.039	0.843
Explanation (E)[b]	0.035	1	0.035	0.012	0.913
P × E	14.394	1	14.394	4.927	0.030
Error	216.209	74	2.922		

Panel C: Planned Contrasts

	Low Explanation	High Explanation	t-value (p-value)[c]
Low Pervasiveness	5.65	6.47	1.42 ($p < 0.10$)
High Pervasiveness	6.59	5.68	1.78 ($p < 0.05$)

[a] Pervasiveness: Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system.

[b] Explanation Detail: In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail.

[c] One-sided.

four treatment groups (all $p < 0.01$). These results indicate that the mean risk assessments are significantly lower for a firm without a control weakness relative to a firm with a control weakness, thereby supporting H1 and findings from archival research.

The analyses we conduct to test H2 and H3 do not include the control group, as control weakness pervasiveness and explanation level cannot be manipulated when no control weaknesses are present. Hypothesis 2 posits that the forward-looking investment risk assessments of nonprofessional investors increase as the pervasiveness of material control weaknesses increases. Panel B

of Table 1 presents the results of an ANOVA model where the dependent variable reflects the participants' assessment of the risk of a moderate stock price decline, and the independent variables are control weakness pervasiveness (low versus high), explanation detail (low versus high), and the interaction term. The coefficient for control weakness pervasiveness is not statistically significant ($F = 0.039$, $p > 0.50$). However, this result must be evaluated cautiously, as there is also a significant interaction between pervasiveness and explanation that we explore further in the analyses that follow.

Hypothesis 3 is tested with the main effect of explanation detail and the interaction of pervasiveness and explanation detail. Results in Table 1, Panel B indicate that explanation detail is not statistically significant ($F = 0.012$, $p > 0.50$), but the interaction of pervasiveness and explanation level is statistically significant ($F = 4.927$, $p = 0.03$).[2] Figure 1 presents the interaction. To further analyze the interaction effect, we conduct two planned contrasts to test the predictions of attribution theory (see Panel C of Table 1). Attribution theory predicts that explanation will increase risk assessments for less pervasive weaknesses and decrease risk assessments for more pervasive weaknesses. A planned contrast comparing low explanation (5.65) to high explanation (6.47) within the low-pervasiveness condition indicates that assessed risk is significantly higher for higher levels of explanation detail, relative to lower levels of explanation detail ($t = 1.42$, $p < 0.10$, one-tailed test). A second planned contrast comparing low explanation (6.59) to high explanation (5.68) within the high-pervasiveness condition indicates that assessed risk is significantly higher for low explanation, relative to high explanation ($t = 1.78$, $p < 0.05$, one-tailed test).

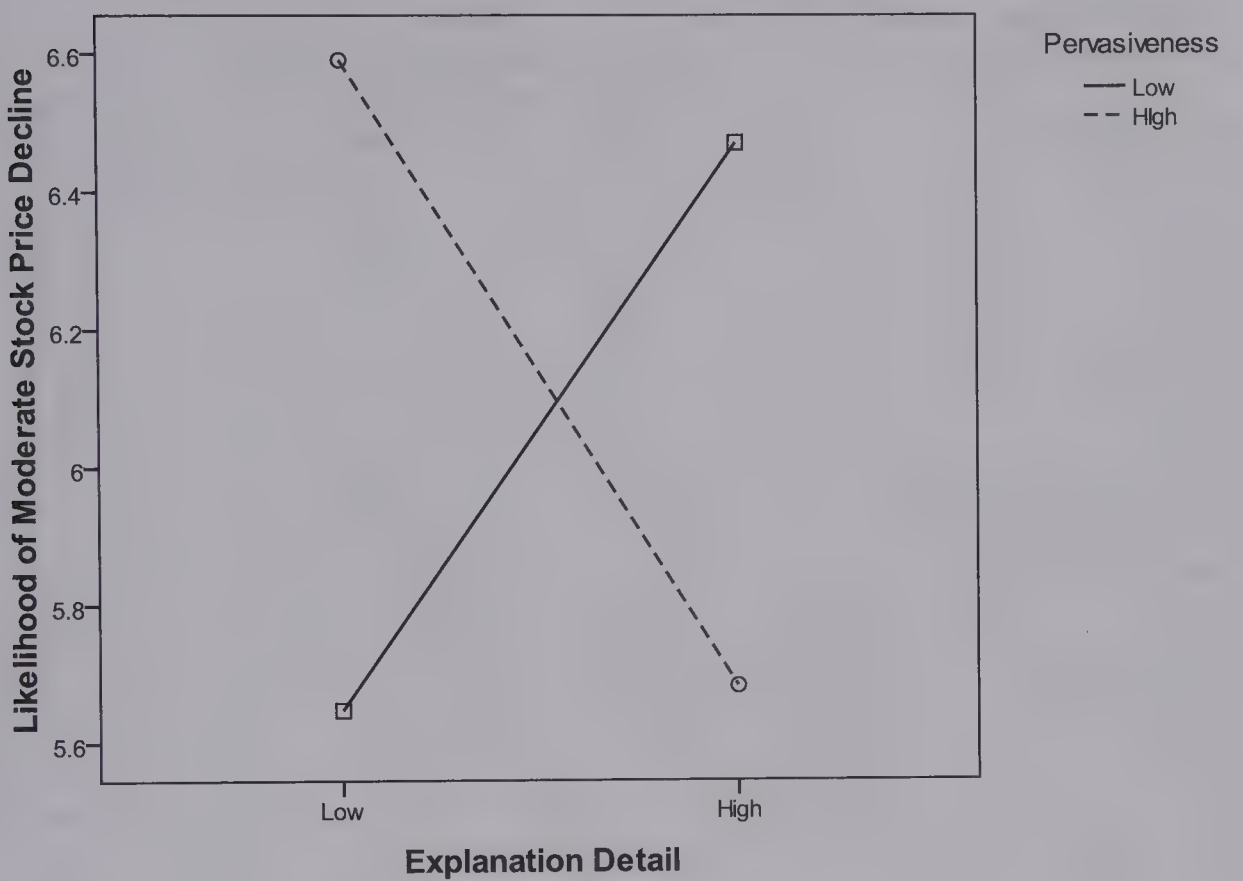
The pattern in Figure 1, the planned contrast results, and the mean responses all support an attribution theory perspective. Contrary to the expectations of corporate directors, detailed explanations of control weakness pervasiveness increase risk assessments for isolated control weaknesses, but decrease risk assessments for pervasive control weaknesses. This finding is further analyzed and discussed in the "Supplemental Analyses and Additional Data Collection" section.

Supplemental Analyses and Additional Data Collection

Attribution theory suggests that trust in management is affected by differences in disclosure expectations versus actual disclosures. Following attribution theory, we expect that when management describes the isolated nature of less pervasive material weaknesses, investors' expectations that management will attempt to cover up bad news will be confirmed. Thus, trust in management will decline. Similarly, when investors are surprised by management's detailed disclosures that emphasize the negative nature of a pervasive weakness, attribution theory predicts that trust in management will increase. Table 2 reports the results of an ANOVA model where the dependent variable is trust in management and the independent variables are control weakness pervasiveness, explanation detail, and the interaction term. The analysis indicates that perceptions of trust display a similar pattern of results to perceptions of changes in stock price. There is a statistically significant interaction of pervasiveness and explanation detail ($p = 0.005$). Figure 2 depicts the interaction. Increasing explanation detail for pervasive weaknesses increases trust in management, while increasing explanation detail for isolated weaknesses decreases trust in management. This result is supported by planned contrasts that test for attribution theory's predictions for differences in trust perceptions across levels of explanation (see Panel C, Table 2). A planned contrast comparing low explanation (5.76) to high explanation (5.06) within the low-pervasiveness condition indicates that trust in management is significantly lower for higher levels of explanation detail, relative to lower levels of explanation detail ($t = 1.43$, $p < 0.10$, one-tailed test). A second

[2] We repeated the analyses described above with our alternative measure of the effects of the control weakness disclosure on stock price. The hypotheses tests produce the same statistical conclusions.

FIGURE 1
Effects of the Interaction of Pervasiveness and Explanation Detail on Assessments of the Likelihood of a Moderate Stock Price Decline



The figure displays the estimated marginal means from the ANOVA analysis of participants' ratings of the likelihood of a moderate stock price decline.
Pervasiveness: Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system.
Explanation Detail: In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail.

planned contrast comparing low explanation (5.45) to high explanation (6.64) within the high-pervasiveness condition indicates that trust in management is significantly higher for higher levels of explanation detail, relative to lower levels of explanation detail ($t = 2.72$, $p < 0.01$, one-tailed test).

To further examine the trust findings, we employ path analysis (Figure 3). The relationship between the interaction of pervasiveness and explanation level and trust in management is positive and statistically significant ($\beta = 0.327$, $p < 0.01$), and the relationship between trust and assessments of risk is negative and statistically significant ($\beta = -0.223$, $p < 0.05$). In addition, the interaction of pervasiveness and explanation detail has a direct, negative effect on assessments of investment risk ($\beta = -0.567$, $p < 0.05$). From the model, it is apparent that the interaction of

TABLE 2

Descriptive Statistics and ANOVA
(Dependent Variable = Trust in Management)

Panel A: Descriptive Statistics
Mean (Standard Deviation) {Sample Size} Responses across Treatment Conditions

Pervasiveness[a]	Explanation Detail[b] Low	Explanation Detail[b] High	Main Effect Pervasiveness
Low	5.76	5.06	5.41
	(1.60)	(1.30)	(1.47)
	{17}	{17}	{34}
High	5.45	6.64	6.05
	(1.10)	(1.71)	(1.54)
	{22}	{22}	{44}
Main Effect: Explanation	5.59	5.94	5.77
	(1.33)	(1.72)	(1.54)
	{39}	{39}	{78}

Control Group: 6.86 (1.34) {19}

Panel B: ANOVA

Source	Sum of Squares	df	Mean Square	F	p-value
Pervasiveness (P)[a]	7.702	1	7.702	3.688	0.059
Explanation (E)[b]	1.086	1	1.086	0.520	0.473
P × E	17.086	1	17.086	8.181	0.005
Error	154.550	74	2.088		

Panel C: Planned Contrasts

	Low Explanation	High Explanation	t-value (p-value)[c]
Low Pervasiveness	5.76	5.06	1.43 ($p < 0.10$)
High Pervasiveness	5.45	6.64	2.72 ($p < 0.01$)

[a] Pervasiveness: Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system.

[b] Explanation Detail: In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail.

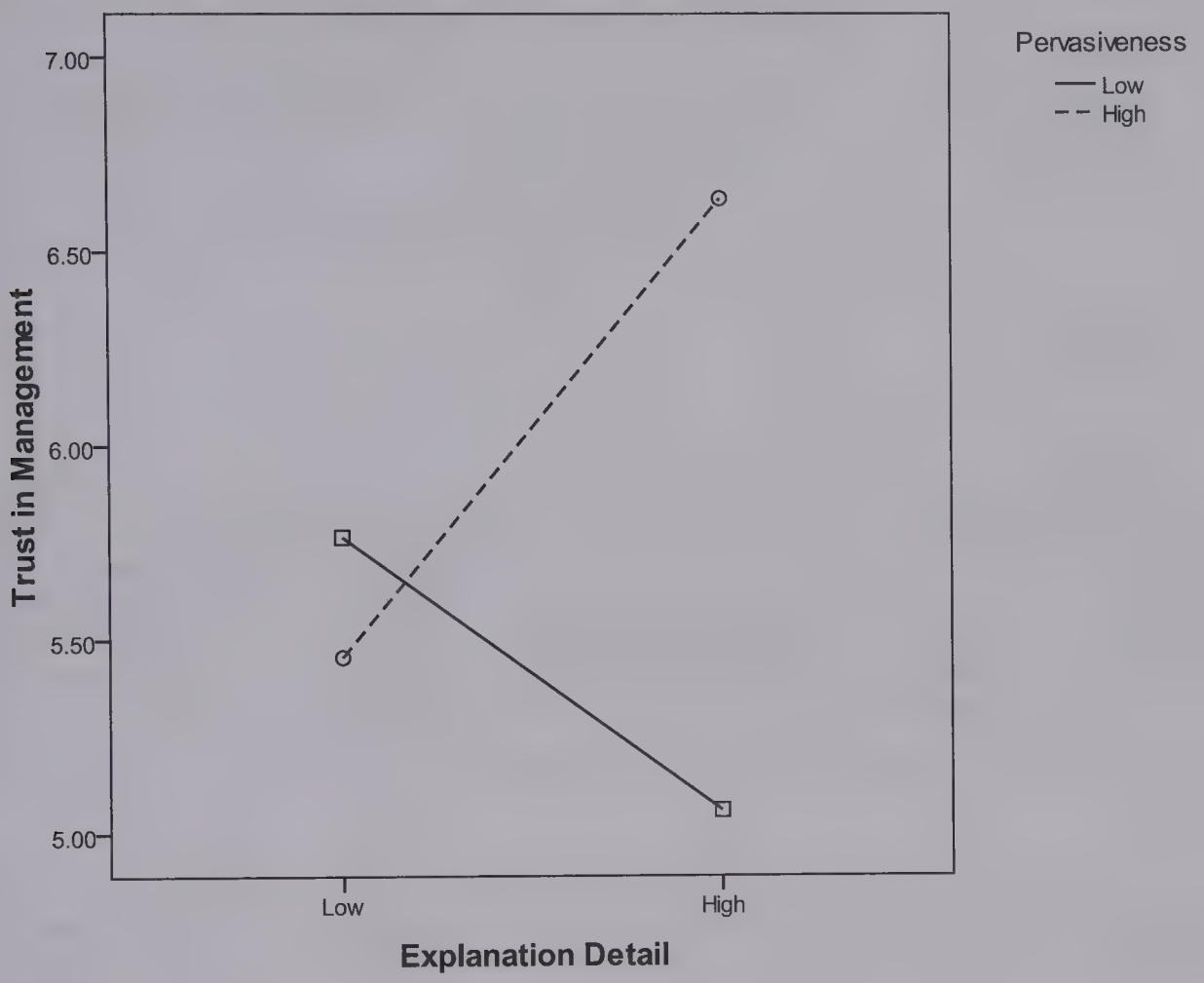
[c] One-sided.

pervasiveness and explanation detail indirectly affects risk assessments through effects on trust in management, but the interaction also directly influences risk assessments. Thus, trust explains a portion of the effects of the interaction on risk assessments, but not all of the effects. There may be other mediating variables that also explain the interactive effects of pervasiveness and explanation detail.

Follow-Up Survey

In order to better understand how trust influences perceptions of investment risk, we survey (under controlled laboratory conditions) an additional 47 M.B.A. students with an average of 3.2

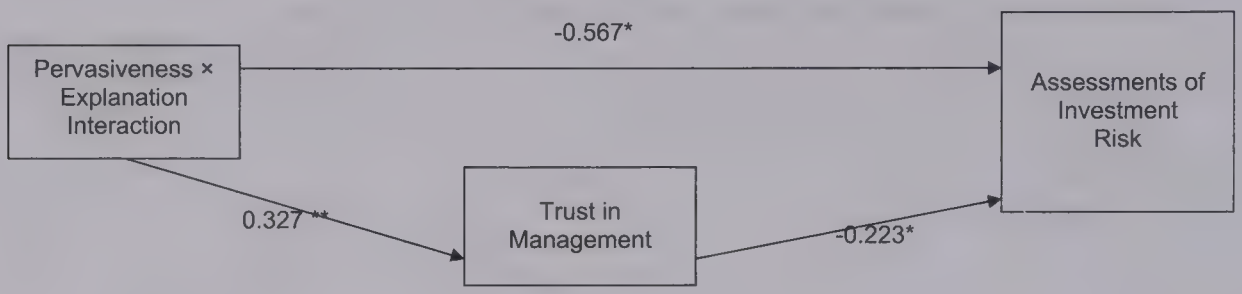
FIGURE 2
Effects of the Interaction of Pervasiveness and Explanation Detail on Trust in Management



The figure displays the estimated marginal means from the ANOVA analysis of participants' ratings of their trust in management.
Pervasiveness: Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system.
Explanation Detail: In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail.

years of management experience. Participants responded to the following three items: (1) "In general, does the trustworthiness of a firm's top management affect the market price of the firm?" [response scale: 1 = Definitely No, 5 = Moderately, 9 = Definitely Yes]; (2) "In general, does the trustworthiness of a firm's top management affect the risk that financial disclosures of the firm will not be accurate?" [response scale: 1 = Definitely No, 5 = Moderately, 9 = Definitely Yes]. (3) "In general, how is the long-term financial performance of a firm affected if management has low levels of trustworthiness?" [response scale: −5 = Very Poor Performance, 1 = No Effect, 5 = Very High Performance]. The mean responses (standard deviations) to survey items are 7.36 (1.69)

FIGURE 3
Structural Model



* $p < 0.05$, two-tailed
** $p < 0.01$, two-tailed

*, ** $p < 0.05$ and $p < 0.01$, respectively, two-tailed.
This figure displays a path analysis that examines whether trust in management mediates the relationship between the interaction of pervasiveness and explanation detail and assessments of investment risk. Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system. In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail. Assessments of investment risk are participants' ratings of the likelihood of a moderate stock price decline. Coefficients represent standardized betas.

for item (1), −3.37 (1.65) for item (2), and 7.28 (1.08) for item (3). For items (1) and (2), the responses are significantly different from a response of "Definitely No" ($p < 0.001$), and are also significantly greater than a response of "Moderately" ($p < 0.001$). For item (3), the mean response is significantly less than "No Effect" ($p < 0.001$). Taken together, the survey responses indicate that concerns about management's trustworthiness affect investors' perceptions of information risk and risk of actual declines in performance. Investors appear to be concerned that, consistent with some recent archival research (Francis et al. 2005; Aboody et al. 2005; Ecker et al. 2006; Kravet and Shevlin 2009), information risk is priced, and they also perceive that management credibility is associated with a firm's future performance. Pilot testing also suggested that trust in management can be associated with concerns about the potential for restatements in the future. While we cannot disentangle which of these perceptions drives our results, it is apparent that the interaction of pervasiveness and explanation detail has the potential to alter perceptions of information risk, risks related to actual future performance, and risks that the market will price management's lack of honest disclosures.

Other Analyses

Next, we perform analyses to rule out alternative interpretations of our findings. Potentially, increased explanation detail could have unintended consequences on investors' perceptions of the pervasiveness of control weaknesses. For example, if additional explanation makes the low-pervasiveness condition appear more pervasive or the high-pervasiveness condition appear less pervasive, then this would explain our interaction results. Table 3 presents an analysis of the effects of pervasiveness and explanation detail on investors' assessments of the pervasiveness of the control weakness. The pattern of means refutes this alternative explanation, and there is a

TABLE 3

Descriptive Statistics and ANOVA
(Dependent Variable = Assessments of Weakness Pervasiveness)

Panel A: Descriptive Statistics
Mean (Standard Deviation) {Sample Size} Responses across Treatment Conditions

	Explanation Detail[b]		
Pervasiveness[a]	**Low**	**High**	**Main Effect Pervasiveness**
Low	5.24	5.12	5.18
	(1.82)	(1.90)	(1.83)
	{17}	{17}	{34}
High	6.18	7.55	6.86
	(1.76)	(1.37)	(1.71)
	{22}	{22}	{44}
Main Effect: Explanation	5.77	6.49	6.13
	(1.83)	(2.01)	(1.94)
	{39}	{39}	{78}

Panel B: ANOVA

Source	**Sum of Squares**	**df**	**Mean Square**	**F**	**p-value**
Pervasiveness (P)[a]	54.595	1	54.595	18.743	0.001
Explanation (E)[b]	7.444	1	7.444	2.556	0.114
P × E	10.521	1	10.521	3.612	0.061
Error	215.551	74	2.913		

[a] Pervasiveness: Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system.

[b] Explanation Detail: In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail.

statistically significant interaction of pervasiveness and explanation detail ($F = 3.612$, $p = 0.06$). Assessments of pervasiveness increase when more explanation is provided (7.55), relative to when less explanation is provided (6.18) for the high-pervasiveness condition. In the low-pervasiveness condition, assessments of pervasiveness do not increase when more explanation is provided (5.12), relative to when less explanation is provided (5.24). Based only upon perceptions of pervasiveness, increased explanation would be expected to decrease assessments of investment risk for less pervasive risks and increase assessments of investment risk for more pervasive risks. We find precisely the opposite results.

Follow-Up Experiment

We conducted a follow-up experiment to provide additional evidence that attribution theory explains our results, rather than the results being an artifact of the wording of our disclosures. The follow-up experiment employed the same dependent and independent variables, and used the same experimental materials, but with one change. Rather than providing participants with the Management Report on Internal Controls and the auditor's opinion on the Management Report on Internal Controls, we informed participants that management had identified a material control weakness and that the auditor concurred that there was a material control weakness. Therefore, there were no

lengthy reports provided in the follow-up experiment. Most importantly, the description of the control weakness was provided by the external auditor, rather than management. The case materials stated the following:[3]

> The external auditor has described the material weakness as follows: "The material weaknesses resulted from a failure to maintain adequate password and access controls over the portion of the Enterprise Resource Planning system that ensures proper recording of sales in the appropriate reporting period."

The purpose of this alternative approach is to remove the effects predicted by attribution theory. While the wording of the disclosures remains the same, the disclosures no longer come from management, and the descriptions of the weaknesses are prepared by the external auditor. Thus, the new experiment removes expectations about management's incentives to downplay bad news and overstate good news because the disclosures are not prepared by management. With management incentives and investor expectations removed, we no longer expect increased explanation detail to increase trust in management for pervasive weaknesses or decrease trust in management for isolated weaknesses. In addition, we no longer expect that increased explanation detail will decrease risk assessments for pervasive weaknesses or increase risk assessments for isolated weaknesses. The follow-up experiment makes it possible to verify that attribution theory explains the interactive effects of pervasiveness and explanation detail on risk assessments.

Tables 4 and 5 present results from the follow-up experiment. Fifty-three M.B.A. students (from the same population of M.B.A. students as the main experiment) with an average of 2.8 years of management experience participated. From Table 4, it is apparent that increased pervasiveness increases assessments of risk (main effect, $F = 27.009$, $p = 0.004$), but there is no longer a significant interaction of pervasiveness and explanation detail ($F = 0.244$, $p > 0.50$). Further, the pattern of means in Panel A of Table 4 and the significant main effect of explanation detail indicate that risk assessments decrease whenever explanations are more detailed ($F = 18.416$, $p = 0.015$). Overall, the findings for assessments of investment risk are consistent with directors' expectations and the tenets of psychological theories of risk (Slovic 1987). That is, pervasiveness increases risk assessments, and additional explanation has the potential to decrease the "unknown" and decrease risk assessments. When management incentives are removed, identical descriptions of the control weaknesses no longer produce the pattern of effects found in the main experiment.

Similarly, analyses of trust in management no longer follow the pattern found in the main experiment. Table 5 presents the trust results. There are no statistically significant effects of pervasiveness or explanation detail on trust in management in the follow-up experiment. Without investors' expectations of management's incentives, and without consistency or inconsistency between these expectations and the content of disclosures from management, the disclosures do not alter trust in management. Taken together, the results of this experiment help to rule out the possibility that the wording of the explanations was the driver of the results. Rather, the expectations that management will downplay bad news and exaggerate good news, and the effects of these expectations on attributions related to management trustworthiness apparently caused the interactive effects of pervasiveness and explanation detail on perceptions of risk.

V. DISCUSSION AND CONCLUSIONS

Investors rely upon audited financial statements to make investment decisions, trusting management's assertions and the independent, expert opinion of the auditors regarding the fairness

[3] This is the statement from the low-pervasiveness/low explanation detail treatment. Each of the four treatment conditions employed descriptions of the control weakness that were identical to the descriptions in the main experiment, and each description was prefaced by the statement that the external auditor was describing the weakness.

TABLE 4

Follow-Up Experiment
Descriptive Statistics and ANOVA
(Dependent Variable = Risk of Moderate Stock Price Decline)

Panel A: Descriptive Statistics

Mean (Standard Deviation) {Sample Size} Responses across Treatment Conditions

Pervasiveness[a]	Explanation Detail[b]		MainEffect Pervasiveness
	Low	High	
Low	5.08	3.67	4.29
	(1.44)	(1.55)	(1.64)
	{12}	{15}	{27}
High	6.29	5.33	5.85
	(1.59)	(2.18)	(1.91)
	{14}	{12}	{26}
Main Effect: Explanation	5.73	4.41	5.06
	(1.62)	(2.01)	(1.93)
	{26}	{27}	{53}

Panel B: ANOVA

Source	Sum of Squares	df	Mean Square	F	p-value
Pervasiveness (P)[a]	27.009	1	27.009	27.009	0.004
Explanation (E)[b]	18.416	1	18.416	18.416	0.015
P × E	0.707	1	0.707	0.244	0.623
Error	141.774	49	2.893		

[a] Pervasiveness: Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system.

[b] Explanation Detail: In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail.

with which the annual financial statements have been presented. However, corporate failures over the past decade suggest that neither the auditors nor management of failed companies were fully committed to protecting the public trust (Powers et al. 2002; Rosen 2003). The Sarbanes-Oxley Act and associated SEC regulations were created to restore investor confidence in financial markets and to provide investors with high-quality information for assessing investment risk. We investigate the effects of control weakness disclosures, as required by Section 404 of the Sarbanes-Oxley Act, on investors' perceptions of investment risk.

The results from laboratory experiments with nonprofessional investors indicate that assessments of investment risk increase for firms with material control weakness disclosures, relative to firms without such disclosures. This finding is consistent with archival research that finds negative stock price reactions resulting from control weakness disclosures, indicating that investors use information about internal control weaknesses when evaluating investment risk.

Directors believe that more detailed explanation of control weaknesses and their pervasiveness will reduce risk assessments when the control weaknesses are not pervasive, and they desire to protect the interests of the firms they represent by clearly explaining the limited effects of less pervasive weaknesses on financial reporting systems. This perspective is supported by theories of

TABLE 5

Follow-Up Experiment
Descriptive Statistics and ANOVA
(Dependent Variable = Trust in Management)

Panel A: Descriptive Statistics
Mean (Standard Deviation) {Sample Size} Responses across Treatment Conditions

	Explanation Detail[b]		
Pervasiveness[a]	**Low**	**High**	**Main Effect Pervasiveness**
Low	4.75	5.13	4.96
	(2.22)	(1.77)	(1.95)
	{12}	{15}	{27}
High	4.79	4.58	4.69
	(1.43)	(2.07)	(1.72)
	{14}	{12}	{26}
Main Effect: Explanation	4.77	4.89	4.83
	(1.80)	(1.89)	(1.83)
	{26}	{27}	{53}

Panel B: ANOVA

Source	Sum of Squares	df	Mean Square	F	p-value
Pervasiveness (P)[a]	0.868	1	0.868	0.248	0.620
Explanation (E)[b]	0.107	1	0.107	0.031	0.862
P × E	1.126	1	1.126	0.322	0.573
Error	171.257	49	3.495		

[a] Pervasiveness: Low pervasiveness involves lack of access controls over a single system, while high pervasiveness involves lack of access controls over the entire system.

[b] Explanation Detail: In the low explanation level, the pervasiveness of the control weaknesses and the specific effects of the control weaknesses are not elaborated. In the high explanation level, the pervasiveness and effects of the control weaknesses are described in detail.

risk perception, but attribution theory predicts different results. We find that investors do recognize the pervasiveness of internal control weaknesses and they use this information to generate investment risk assessments. However, the anticipated benefits of additional disclosure sought by directors are not supported. Rather, results suggest that providing a detailed explanation that a control weakness is isolated and not pervasive increases forward-looking risk assessments relative to not providing such explanation. Conversely, providing an explanation that a control weakness is very pervasive decreases forward-looking risk assessments relative to not providing a detailed explanation. While we find that investors use information about control weakness pervasiveness to assess risk, pervasiveness is a multi-faceted construct (e.g., multiple versus single accounts, larger versus smaller dollar amounts, less versus more vague effects on financial statement). Future research will be necessary to determine which components of pervasiveness are most important to investors, if this is desirable.

We find evidence that changes in trust in management likely explain the pattern we observe. The nonprofessional investors in our experiments believe that management is not trustworthy when detailed explanations are provided for non-pervasive control weaknesses, while investors believe management is trustworthy when detailed explanations are provided for pervasive control

weaknesses. The results indicate that investors perceive that management is being dishonest when detailed explanations are provided for isolated control weaknesses and that management is attempting to cover up the seriousness of the control weakness by claiming that it is non-pervasive. Detailed explanations of pervasive control weaknesses, however, are viewed as indicators that management is trustworthy—that management is openly admitting the serious nature of the control weakness.

Overall, our findings respond to concerns from archival research studies that investors' risk judgments are more complex than suggested by prior accounting risk models and that financial risk judgments are influenced by probabilities, outcomes, and numerous behavioral variables (Koonce et al. 2005). Further, Schwarzkopf (2006) claims that we do not really understand what makes a business situation appear more or less risky to an individual, and therefore, managers should make a greater effort to understand how investors view risk and to consider more open stakeholder dialog, more involved message content, and increased transparency. The explanations provided by management in control weakness disclosure, and the interaction of these explanations with the pervasiveness of the weaknesses have meaningful effects on investors' judgments of risk, and we provide insight into the complex processes that affect these risk judgments.

REFERENCES

Aboody, D., J. Hughes, and J. Liu. 2005. Earnings quality, insider trading, and cost of capital. *Journal of Accounting Research* 43: 651–673.

Baginski, S., J. Hassell, and W. Hillison. 2000. Voluntary causal disclosures: Tendencies and capital market reaction. *Review of Quantitative Finance and Accounting* 15: 371–389.

Beneish, M., M. Billings, and L. Hodder. 2008. Internal control weaknesses and information uncertainty. *The Accounting Review* 83 (3): 665–704.

Das, T., and B. Teng. 2004. The risk-based view of trust: A conceptual framework. *Journal of Business and Psychology* 19 (1): 85–116.

Doss, M. 2004. Section 404 reports on internal control: Impact on ratings will depend on nature of material weaknesses reported. *Moody's Investors Service* (October).

Eagly, A., and S. Chaiken. 1975. An attribution analysis of the effect of communicator characteristics on opinion change: The case of communicator attractiveness. *Journal of Personality and Social Psychology* 32: 136–144.

———, and ———. 1993. *The Psychology of Attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.

Ecker, F., J. Francis, I. Kim, P. Olsson, and K. Schipper. 2006. A returns-based representation of earnings quality. *The Accounting Review* 81: 749–781.

Elliott, W., F. Hodge, S. J. Kennedy, and M. Pronk. 2007. Are M.B.A. students a good proxy for nonprofessional investors? *The Accounting Review* 82 (1): 139–168.

Fischhoff, B. 1995. Risk perception and communication unplugged: Twenty years of progress. *Risk Analysis* 15 (April): 137–145.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39: 295–327.

Hammersley, J., L. Myers, and C. Shakespeare. 2008. Market reactions to the disclosure of internal control weaknesses and to the characteristics of those weaknesses under section 302 of the Sarbanes Oxley Act of 2002. *Review of Accounting Studies* 13 (1): 141–165.

Hirst, D., L. Koonce, and P. Simko. 1995. Investor reactions to financial analysts' research reports. *Journal of Accounting Research* 33: 335–351.

Hodge, R., J. Kennedy, and L. Maines. 2004. Does search-facilitating technology improve the transparency of financial reporting? *The Accounting Review* 79 (3): 687–703.

Hutton, A., G. Miller, and D. Skinner. 2003. The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41: 867–890.

Kasperson, R., R. Ortwin, P. Slovic, H. Brown, J. Emel, R. Goble, J. Kasperson, and S. Ratick. 1988. The social amplification of risk: A conceptual framework. *Risk Analysis* 8: 177–187.

Kerstetter, T., and H. Schulken. 2005. *What Directors Think*. Brentwood, TN: Board Member Inc.

Kinney, W. 1989. Achieved audit risk and the audit outcome space. *Auditing: A Journal of Practice & Theory* (Supplement): 67–84.

Koonce, L., M. McAnally, and M. Mercer. 2005. How do investors judge the risk of financial items? *The Accounting Review* 80 (1): 221–241.

Kravet, T., and T. Shevlin. 2009. Accounting restatements and information risk. *Review of Accounting Studies*.

Loewenstein, G., E. Weber, C. Hsee, and E. Welch. 2001. Risk as feelings. *Psychological Bulletin* 127 (March): 267–286.

Maines, L., and L. McDaniel. 2000. Effects of comprehensive income characteristics on nonprofessional investors' judgments: The role of financial statement presentation format. *The Accounting Review* 75 (2): 179–207.

Mercer, M. 2005. The fleeting effects of disclosure forthcomingness on management's reporting credibility. *The Accounting Review* 80 (2): 723–744.

Messier, W., and L. Austen. 2000. Inherent risk and control risk assessments: Evidence on the effects of pervasive and specific risk factors. *Auditing: A Journal of Practice & Theory* 19 (2): 119–131.

Ogneva, M., K. Raghunandan, and K. Subramanyam. 2007. Internal control weakness and cost of equity: Evidence from SOX Section 404 certifications. *The Accounting Review* 82 (5): 1255–1298.

Olsen, R. 1997. Investment risk: The experts' perspective. *Financial Analysts' Journal* 53 (March/April): 62–66.

Orenstein, E. 2004. Sarbanes-Oxley Section 404: Remediation, communication, education. *Financial Executive* 20 (8): 28–30.

Palmrose, Z.-V., V. Richardson, and S. Scholz. 2004. Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics* 37: 59–89.

Pidgeon, N., R. Kasperson, and P. Slovic, eds. 2002. *The Social Amplification of Risk*. London, U.K.: Cambridge University Press.

Powers, W., R. Troubh, and H. Winokur. 2002. *Report of the Investigation by the Special Investigative Committee of the Board of Directors of Enron Corp*. Austin, TX: Enron.

PricewaterhouseCoopers (PWC). 2005. *What Directors Think*. New York, NY: PWC.

Public Company Accounting Oversight Board (PCAOB). 2007. *An Audit of Internal Control Over Financial Reporting That Is Integrated with An Audit of Financial Statements*. Auditing Standard No. 5. Available at: http://pcaobus.org/Standards/Auditing/Pages/Auditing_Standard_5.aspx.

Rosen, R. 2003. Risk management and corporate governance: The case of Enron. *Connecticut Law Review* 35: 1157–1184.

Ryan, L., and A. Buchholtz. 2001. Trust, risk, and shareholder decision making: An investor perspective on corporate governance. *Business Ethics Quarterly* 11 (1): 177–201.

Schwarzkopf, D. 2006. Stakeholder perspectives and business risk perception. *Journal of Business Ethics* 64 (4): 327–342.

Securities and Exchange Commission (SEC). 2003. *SEC Implements Internal Control Provisions of Sarbanes-Oxley Act; Adopts Investment Company R&D Safe Harbor*. Available at: http://www.sec.gov/news/press/2003-66.htm.

Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32: 38–60.

Slovic, P. 1987. Perception of risk. *Science* 236 (April): 280–285.

Soffer, L., S. Thiagarajan, and B. Walther. 2000. Earnings preannouncement strategies. *Review of Accounting Studies* 5: 5–26.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107–204 [H.R. 3763]. Washington, D.C.: Government Printing Office.

Weber, E. 1988. A descriptive measure of risk. *Acta Psychologica* 69 (November): 185–203.

Weber, E., and W. Bottom. 1989. Axiomatic measures of perceived risk: Some tests and extensions. *Journal of Behavioral Decision Making* 2 (April-June): 113–131.

———, and ———. 1990. An empirical evaluation of the transitivity, monotonicity, accounting, and conjoint axioms for perceived risk. *Organizational Behavior and Human Decision Processes* 45 (April): 253–275.

THE ACCOUNTING REVIEW
Vol. 85, No. 5
2010
pp. 1809–1822

American Accounting Association
DOI: 10.2308/accr.2010.85.5.1809

BOOK REVIEWS

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Business, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

PETER O. CHRISTENSEN and GERALD A. FELTHAM, *Equity Valuation* (Hanover, MA: Foundations and Trends® in Accounting, 2009, ISBN 978-1-60198-272-8, Vol. 4, No. 1, pp. x, 115).

Let me begin by saying that this is a great booklet.[1] The authors propose using the riskless spot rate and forward rates, as opposed to risk-adjusted discount rates, in discounting future earnings and cash flows to value equity. It is a key conceptual departure from practitioners' orthodoxy. Like most analytical works, the beauty of its conciseness and precision is evident throughout. The authors also put together various parts of the analysis through a consistent and overarching theme, making sure that readers are not lost in the mathematical details and do not lose sight of the big picture. In my mind, they strike an almost perfect balance between making the book precise and solid for academics, and at the same time engaging and persuasive for practitioners. Because of this, it appeals to a wide range of readers. I highly recommend this book to anyone who is interested in the area of accounting-based valuation. Overall, on the scale of 1 to 5, I would rate this book a solid 4.

Although the authors did a great job in making this book engaging and accessible, it is definitely not a beginners' book. Readers are assumed to possess a fair amount of knowledge about the valuation literature as well as the common modeling ingredients and techniques. The good news is that the authors make frequent references to a couple of related books with more detailed analysis (Christensen and Feltham 2003; Penman 2007). Thus, for those who are interested in delving into more details, these are readily available in those books.

The book is titled, broadly, as *Equity Valuation*. It deals mostly with one particular issue: In estimating equity value through discounting future payoffs, should we adjust for risk in the numerator (i.e., the forecasted future payoffs) or in the denominator (i.e., the discount rate)? This is an important question. In performing accounting-based valuation analysis, one usually replaces cash flow with either residual income or abnormal earnings growth. Either way we need to deal with the issue of what discount rate should be used in calculating residual income. Following the traditional approach of incorporating risk into such discount factors poses two serious problems. First, when such discount factors are stochastic, which is often the case with the non-deterministic term structure of interest rates, inflation, and/or risk adjustment, we are faced with the difficult task of determining the covariance between the discount rate and the book value in calculating further residual earnings. Second, and more importantly, if the discount rate already incorporates risk adjustment, then it contradicts one of the important goals of the whole analysis, namely, to use accounting data as input for the calculation of proper risk adjustment. This is why it is crucial for us to establish a theory of how to adjust for risk in the numerator, leaving the discount rate

[1] I thank Charles Lee and Jim Ohlson for helpful comments and suggestions.

deterministic. Of course, even putting aside the need for incorporating accounting in equity valuation, it is utterly essential to be able to determine a proper discount rate in the valuation of any claim.

Within the equity valuation literature, this issue has been largely ignored. Since the establishment of the CAPM model, the finance literature on valuation has mostly flourished in some other areas, most noticeably market micro-structure, and most recently behavioral finance. Within the neoclassical framework, work has been done to address issues such as adaptive learning, stochastic expected stock return, and stochastic risk premium. This paper revisits a quite fundamental issue, i.e., should we adjust risk in the numerator or the denominator? It extends the recent works by Feltham and Ohlson (1999) and Christensen and Feltham (2003), which are direct descendants of the seminal work of Rubinstein (1976), who characterized the relationship between the valuation of cash flow streams and the prices of zero-coupon bonds in a multi-period setting under uncertainty.

The authors advocate adjusting risk in the numerator rather than in the denominator. In other words, they propose using the riskless spot rate and forward rates as the discount rates in valuation. The authors list several problems with the traditional approach in estimating the risk-adjusted discount rate. They are:

- many of the ideas are developed in a single-period setting, which cannot be easily carried over to multi-period setting. Most noticeably, calculating cost of capital as a weighted average of cost of equity and cost of debt is likely to introduce error into a multi-period setting, because the internal rate of return on a portfolio of assets differs from the weighted average of the internal rates of returns on the individual assets;
- with stochastic expected stock returns, the estimate of average return becomes a biased measure of the average cost of capital used in valuation, due to Jensen's inequality; and
- even with recent asset-pricing models, which address the stochastic expected return issue with date-specific, risk-adjusted discount rates, such rates may not exist in certain cases.

The authors develop an equilibrium asset-pricing model with risk adjustments depending on the time-series properties of cash flows and the accounting policy. They show that operating characters such as the growth and persistence of earnings can affect the risk adjustment.

What are the highlights of this book? The book contains five chapters and three appendices. Chapters 2 to 5 each contain separate yet closely related topics. Chapter 2 reviews and identifies problems with the implementation of the classical model. In Chapters 3 to 5, the authors develop an accounting-based, multi-period asset-pricing model with HARA utility. My preferences are Chapters 2 and 5. Chapter 2 contains a critical review of the classical valuation approach with a constant risk-adjusted discount rate. As noted above, the authors highlight several problems in estimating these models. Many of these issues are not properly acknowledged and/or dealt with in many of the textbooks. The authors provide a nice step-by-step analysis of the problems and possible solutions.

Chapter 5 contains the punch line. The authors push ahead with the idea of adjusting risk in the numerator, and deal with the thorny issue of identifying and simplifying the so-called "pricing kernel." Although the final model involves a rather simplifying assumption of a simple VAR model of the stochastic processes of residual income and for the consumption index, it provides striking and promising ideas of how to estimate and adjust for risk based on fundamentals, as opposed to stock return. It provides a nice illustration of how to incorporate time-change risk characteristics of firms with the change in firms' operations (captured by the change in residual income). This is very encouraging.

There are some unsettling issues in this book. Not surprisingly, I find the authors' review of the classical valuation approach to be somewhat tilted toward the negative side. For instance, many of the problems cited arise from the practice of estimating a single, constant risk-adjusted discount rate for all future periods. This seems to be based on the assumption that firms' risk characteristics do not change materially over future periods. Of course, this is a grossly simplified approach in dealing with the issues of time-changing interest rates and inflation. To me, errors introduced by such an approach reflect more the shortcomings in the empirical or practical implementation, rather than the shortcomings in the valuation approach *per se*. As noted by the authors, using date-specific discount rates can avoid many of the problems. After all, under most circumstances in a neo-classical framework, putting the risk adjustment in the numerator or in the denominator may simply be an easy mathematical transformation. In some cases, of course, adjusting risk in the denominator does not lead to any solution to the problem. In that sense, adjusting in the numerator is more flexible.

After finishing the book, I asked myself the following question: Am I convinced that the practice of adjusting risk in the discount rate should be abolished? The answer seems unclear, for a couple of reasons. First, despite the authors' admirable effort in bringing context to it, the concept of "consumption index" still seems rather elusive. As a result, it lacks the appeal of the traditional CAPM, namely, a clear and intuitive idea of risk adjustment.

Second, given that the "consumption index" or pricing kernels are by nature event-price measures, measuring the "consumption index" empirically is likely to introduce considerable error. Such error, in all likelihood, may exceed what we observe with the calculation of WACC using CAPM. Of course, this is ultimately an empirical issue. From a difference perspective, such concerns also make the issue more interesting and engaging for future research.

In sum, despite the somewhat unsettling issues mentioned above, this work makes a great handbook for anyone who is interested in pursuing research in accounting-based equity valuation. Since each chapter essentially deals with one independent, yet interrelated issue, the work constitutes a nice reference book. It is particularly useful for doctoral students who are interested in pursuing work in this area. It would be a fun and engaging exercise for them to fill in the missing proofs in each step of the derivations. I highly recommend it.

Traditionally, core courses in finance and accounting divide materials roughly along the line of numerator and denominator: accounting professors teach students how to analyze and forecast future earnings and cash flows, and finance professors teach students how to properly discount them. Given this division of labor, I doubt that my finance colleagues would happily embrace the idea of discounting everything with zero-couple rate since that does not leave them much to talk about. But as an accountant, I like that idea.

REFERENCES

Christensen, P. O., and G. A. Feltham. 2003. *Economics of Accounting: Volume I—Information in Markets*. New York, NY: Springer.

Feltham, G. A., and J. A. Ohlson. 1999. Residual earnings valuation with risk and stochastic interest rates. *The Accounting Review* 74 (2): 165–183.

Penman, S. 2007. *Financial Statement Analysis and Security Valuation*. New York, NY: McGraw-Hill.

Rubinstein, M. 1976. The valuation of uncertain income streams and the pricing of options. *The Bell Journal of Economics* 7 (2): 407–425.

XIAO-JUN ZHANG
Associate Professor of Accounting
University of California, Berkeley

PAUL M. CLIKEMAN, *Called to Account: Fourteen Financial Frauds that Shaped the American Accounting Profession* (New York, NY: Routledge, 2009, ISBN 978-0-415-99698-3, pp. x, 345).

This book is a complete package for auditing professors and students because it includes more than just descriptions of 14 significant accounting scandals that occurred in the United States. The book is divided into seven parts. Each part begins with a chapter that provides a context for the discussion of the frauds described in that part and concludes with a chapter describing the regulatory fallout from those frauds. For the instructor, an appendix includes a set of discussion questions for each chapter and a table integrating *Called to Account* with popular auditing texts. Both instructor and student will find the extensive lists of bibliographic references should they wish to refer to original sources.

Called to Account does an excellent job of presenting the frauds as catalysts for the changes that have taken place in regulation, in accounting and auditing standards, and in accounting practice over the period from the early 1900s to the first part of the 21st century. The information about the frauds is presented in enough detail so that the reader can see why the fraud was successful. I recall Professor John Willingham saying a number of years ago that most partners and managers in public practice had not been involved in an audit where there had been a fraud and thus were not as well equipped as they might be to recognize a fraud at an audit client. *Called to Account* provides enough detail that the readers can easily imagine themselves in the picture and being required to make the decisions the auditor had to make.

Chapter 1, "Scandal and Reform," provides a history of public accounting and regulation in America beginning with the public's interest in the stock market in the early years of the 20th century through to the collapse of WorldCom in 2002.

Part I Birth of a Profession

Chapter 2, "Out of Darkness," examines the early history of capital markets and the public accounting profession. The author states, "U.S. Steel defied convention when it published its first set of audited financial statements in 1903" (p. 17). Prior to that date most companies believed that opaqueness was preferred (and cheaper). The chapter describes the early attempts by the fledgling public accounting profession to encourage better financial reporting.

In Chapter 5, "Into the Spotlight," Clikeman ties the Ivar Kreuger and McKesson & Robbins frauds, discussed in Chapters 3 and 4, to the Securities Act of 1933 and the Securities Exchange Act of 1934 and the setting up of the Securities and Exchange Commission (SEC). He also describes the development of Statement on Auditing Procedure No.1, *Extensions of Auditing Practice*, by a committee of the American Institute of Accountants (AIA). This standard mandated the auditor's observation of client inventory counts and direct confirmation of accounts receivable.

Part II The Profession's Principle Problem

Chapter 6, "Generally Accepted Accounting Principles," traces the development of generally accepted accounting principles (GAAP) from Pacioli in 1494 through the establishment by the AIA's Committee on Accounting Procedure (CAP) in 1938 and the Accounting Principles Board (APB) in 1959. Clikeman chronicles the problems that the CAP and the APB had in persuading the public accounting profession and business community of the need for better accounting.

The frauds discussed in this part are National Student Marketing and Equity Funding.

Chapter 9, "Déjà vu," is a continuation of Chapter 6. The formation of the Financial Accounting Standards Board (FASB) is described. The discussions in Chapter 6 about the problems that the CAP and APB had and in this chapter the problems that the FASB had in setting accounting principles that were generally accepted provide a clear picture of how difficult it has been to build consensus. The chapter also includes a discussion of Congress' Moss and Metcalf investigations and the Cohen Commission on Auditors' Responsibilities (1978), which stated in its report that an "expectation gap" existed between what auditors believed they should do and what the public expected of auditors.

Part III The Savings and Loan Crisis

Chapter 10, "It's a Wonderful Life," discusses the Garn-St. Germain Bill passed by Congress in 1982, which "greatly expanded the types of assets savings and loans were permitted to own" (p. 86). Clikeman describes how savings and loan accounting, including Regulatory Accounting Principles (RAP), allowed the savings and loan associations to manipulate their earnings to appear to be more profitable than they were. He also describes how the Federal Savings and Loan Insurance Corporation lowered the savings and loan net worth requirement in the 1980s from 5 percent to 4 percent to 3 percent. Moral hazard became a problem.

The frauds discussed are ESM Government Securities and Lincoln Savings & Loan.

Chapter 13, "Bank Robbers," describes Congressional hearings that were critical of the management practices of the savings and loans and also of the auditors of those savings and loans. In 1989 and 1991, the Senate and House passed two acts dealing with problems with savings and loans and with banks, respectively. Both acts added significant regulation for both kinds of financial institutions. The chapter includes a discussion of the Foreign Corrupt Practices Act (FCPA) passed in 1977 and the battles between the SEC and the public accountants and industry over reporting by management on internal control.

Part IV The Expectation Gap

Chapter 14, "Auditors and Fraud," is a concise history of auditing from 3200 B.C. to the beginning of the 20th century, followed by a discussion of the evolution of the auditor's responsibility for the detection of fraud from the late 1800s to the mid-1980s.

The two frauds reviewed are ZZZZ Best and Crazy Eddie.

Chapter 17, "Closing the Gap," describes the creation in 1985 of the National Commission on Fraudulent Financial Reporting (Treadway Commission) by the Committee of Sponsoring Organizations (COSO). A conclusion of the Treadway Commission's report, which was issued in 1987, was that management bore the primary responsibility for financial reporting. The chapter also discusses the nine "Expectation Gap" auditing standards and the replacement of Statement on Auditing Standards (SAS) No. 53 with SAS No. 82, which in turn was replaced by SAS No. 99.

Part V The Litigation Crisis

Chapter 18, "Auditors' Legal Liability," begins with the statement, "Auditors in the early 1900s were subject to scant discipline" (p.159), and then provides a brief summary of auditors' liability to third parties. The rest of the chapter provides a historical perspective of auditors' liability to clients, auditors' liability to third parties, and auditors' liability under statutory law. I found this to be the weakest chapter in the book, because the coverage of the topic is too brief to do justice to the topic; auditing texts do a better job. Chapter 21 does provide more extensive coverage of the legal environment. I believe the two chapters might have been combined.

The two cases discussed in this part (Fund of Funds and MiniScribe) are very different, although both relate to litigation against the auditor.

The Fund of Funds case discusses the dilemma that Arthur Andersen & Co. found itself in when the same office, and the same partner, audited two clients (King Resources and Fund of Funds) that had transactions between them. Client confidentiality was not accepted as a defense. It would have been instructive if the author, perhaps in a sidebar, had referred to the case of *Waggenheim (Consolidata Services, Inc.) v. Alexander Grant & Co.*, 19 Ohio App 3rd 7,482 N.E. 2d 955 (1982), where the accounting firm of Alexander Grant & Company lost the case because the court ruled the firm had violated confidentiality.

Chapter 21, "Litigation Reform," chronicles the significant concern that arose from the settlements and from the reported "6,000 liability suits seeking more than $15 billion in damages" (p. 184) from U.S. accounting firms in 1992. The author describes the accounting profession's perception of the causes of the problem: joint and several liability and the inability of a successful defendant to collect costs, and extreme punitive damages that were being awarded.

The Private Securities Litigation Reform Act of 1995 "established proportionate liability for defendants who did not knowingly engage in securities fraud" (p. 187). In the mid-1990s, state legislation was amended to permit CPA firms to practice as limited-liability partnerships.

Part VI Beginning of the End

Chapter 22, "Auditor Independence," provides a thorough discussion of the topic, beginning with a historical perspective. The issues relating to independence are discussed: auditors are paid by their clients, accounting firm staff being hired by clients, nonaudit services, and joint ventures with clients. The author discusses the creation of the Independence Standards Board created by the American Institute of Certified Public Accountants. One of the messages that the chapter conveys is that the CPA profession and especially the largest firms fought any attempt to regulate them and constrain their activities.

Waste Management and Sunbeam are discussed in this part.

Chapter 25, "End of the Millennium," continues the discussion of independence that began in Chapter 22. There is a lengthy discussion of the SEC's proposed and final rules on independence issued in 2000. The author states that the SEC mandated fee disclosure, which resulted in a number of the larger corporations restricting the non-audit services purchased from their auditors. The chapter continues with a discussion of the SEC's Staff Accounting Bulletin No. 99 and the requirement that auditors also consider qualitative materiality.

Part VII From Profession to Regulated Industry

Chapter 26, "Professionalism," traces the evolution of public accounting from early in the 20th century, when it was truly a profession, through the gradual erosion of its standards and business model into a full-service business. The author uses the history of Arthur Andersen to demonstrate that evolution.

The two frauds that probably had the most impact on the accounting profession, Enron and WorldCom, are the focus of this part.

Chapter 29, "The Perfect Storm," discusses the fall of Arthur Andersen as a precursor for the development and passing of the Sarbanes-Oxley Act of 2002 (SOX). There is a discussion of both the Act and its implementation. There is also an examination of SAS No. 99 and its expanded requirements for auditors to consider with respect to possible fraud. The author recounts the criticism of SOX's Section 404 and its costs and impact on reporting firms in terms of the cost in dollars and time. The author states, "Unfortunately, nobody has yet documented a discernible improvement in earnings quality post-Sarbanes-Oxley" (p. 278). I can only conclude that the book was completed before Ashbaugh-Skaife et al. (2008, 2009) were published. And I assume that completion of the book predates the release of Auditing Standard No. 5 by the Public Company Accounting Oversight Board.

Chapter 30, "Conclusion," summarizes the preceding 29 chapters with respect to the development of auditing and audit standards, the development of accounting standards, the changing role of auditor responsibility for fraud, and the evolution of public accounting firms and regulation, and is a fitting conclusion to the book.

There are books published that include cases on fraud (*Fraud Casebook: Lessons from the Bad Side of Business* [Wells 2007] and *Contemporary Auditing: Real Issues & Cases, Update* [Knapp 2010]), but I believe that their objectives are different, and so they do not include the rich discussion of the context and impact of the frauds on accounting and auditing standards, regulation, and the law that are included in *Called to Account: Fourteen Financial Frauds that Shaped the American Accounting Profession.*

REFERENCES

Ashbaugh-Skaife, H., D. Collins, W. Kinney, and R. LaFond. 2008. The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review* 83 (1): 217–250.

———, ———, ———, and ———. 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity capital. *Journal of Accounting Research* 47 (1): 1–43.

Knapp, M. C. 2010. *Contemporary Auditing: Real Issues & Cases, Update*. Cincinnati, OH: South-Western.

Wells, J. T. 2007. *Fraud Casebook: Lessons from the Bad Side of Business*. Association of Certified Fraud Examiners. Hoboken, NJ: Wiley.

W. MORLEY LEMON
Professor *Emeritus* of Accountancy
University of Waterloo

E. RICHARD CRISCIONE, *Abraham J. (Abe) Briloff: A Biography*, Studies in the Development of Accounting Thought, Volume 11 (Bingley, U.K.: JAI/Emerald Group, 2009, ISBN: 978-1-84855-588-4, pp. xxi, 258).

There are not many individuals with deep roots in the academic accounting profession whose biography would have the potential to be both lively and have content enough for a 250+ page book. Most accounting academics are simply not at the center of events that have changed the accounting profession or the broader business community. The vast majority of accounting academics with extensive interaction with the accounting profession (such as those who serve on accounting rule-making bodies) have worked within the system in a relatively non-dramatic and non-confrontational way; they have chosen not to want to get in front of the parade or not to lead the band to make major change. Sandy Burton, who was a proactive Chief Accountant of the Securities and Exchange Commission in the 1970s, is perhaps an important exception.

Abraham Briloff is neither quiet nor lacking in his willingness to be confrontational. Indeed, his critics have often taken the route of making personal attacks on his willingness to be at the center of controversy. His flowery language has only added to his colorful stature. It is difficult to think of many accounting academics whose professional life over multiple decades (especially 1965 to about 2000) would make engaging reading. Criscione's book (drawn from his Ph.D. dissertation at The University of Mississippi) is a commendable work of bringing together a vast amount of diverse material in a very instructive and interesting way. While I spent much time years ago going through Briloff's writings, I still found much that was both new and interesting in this lively book. It is well written with a solid underpinning in a broad base of field research.

The beginning of my own interaction with Briloff was while lecturing in the downtown night program of The University of Chicago. A senior executive of a company whose accounting practices had been criticized by Briloff in *Barron's* challenged me to explain why the stock price of his company plummeted at the time the article came out. The executive asserted that the article was both "profoundly wrong and offensive" as well as being "based only on a reading of publicly available information." My response during the class did not satisfy the executive, and he challenged me in front of the class to come back next week with a "better story." In the seven days before the next class, I feverishly collected stock market and other information on as many Briloff articles as I could access. I developed exhibits that had patterns very similar to what was subsequently published in a *Journal of Accounting Research* (Foster 1979) article. Nick Dopuch, *JAR*'s very influential editor at the time, told me that, while my "Briloff and the Capital Market" paper was not a "normal research piece," it was a "fun read" and correctly predicted it would get a lot of attention. I suspect that I was chosen as the reviewer for this book because of that 1979 paper and a related paper in 1987 (Foster 1987). Those were published a long time ago. However, it was a joy to take the time many years later to read this book and reflect on Abe Briloff and his writings.

Criscione's choice of a biography subject was an excellent one. First, Briloff had a long career in which he exposited some important themes regarding "actual and perceived independence of the accounting professional" and the "attempts by management to use accounting choices to present an overly rosy picture of their company's fundamentals." These are important and broadly understandable issues in both academe and the general business

community. His themes were not topics of interest only to a narrow inner circle of academics who had little interest in whether their writings had any relevance in the commercial world. Second, Briloff wrote and spoke with a passion and in colorful phrases that resulted in much polarization of opinion about his ideas and him personally. This adds to the narrative in Criscione's book being very readable. Briloff was hard-hitting in his criticisms, and the responses to him likewise were far from timid. This makes for fun reading, although at times the tone (on both sides) makes the reader cringe. His SafeCard Services article in *Barron's* led to a company press release that used Briloff's legal blindness as the basis for a low, cheap shot at him: "Sadly for the 'Professor,' he is legally blind, and, therefore, not under any obligation it seems to determine the facts" (quoted on page 34 of book). Third, several of the positions Briloff took with regard to the accounting profession—such as that management consulting and auditing should *not* be done by the same firm—were subsequently adopted by regulators and politicians.

Briloff was able to use several high-profile public platforms to propound his ideas. It is hard to think of other advocates of major change in the accounting profession who had such effective "bully pulpits" to present their case on an ongoing basis over many years. Criscione describes how Briloff started writing for *Barron's*. Briloff wrote a response to an article in *The Accounting Review* that criticized his position on pooling accounting, but the *Review* rejected his paper. He showed his "rejected response" to *Barron's*, which decided to run with it under the title of "Dirty Pooling" (pp. 28–29). This was the first of many *Barron's* articles by Briloff over an extended time period. Over the years, he had the very strong support of the then editor of *Barron's*, Robert M Bleiberg. Multiple lawsuits were brought against both Briloff and *Barron's*, asserting, among other things, "libel," "defamation," and "negligence." The second major platform that especially irked many in the inner circle of the accounting profession was his frequent appearances at Washington Congressional hearings. Politicians who sought very strongly worded and argued positions criticizing key aspects of the "cozy world of the auditing/tax/advisory" triangle of most accounting firms knew that Briloff would provide very targeted positions as well as colorful "sound bites." The fact that academic defenders of the accounting profession (such as Robert N. Anthony of Harvard Business School) were willing to attack Briloff at the personal level as well as for his arguments only added to the debate being viewed by many as "high theatre." I found Criscione's Chapter 6 on "Mr. Briloff Goes to Washington" an especially well-presented one in an area that surprised me with regard to how much both politicians and Briloff were able to effectively "use" each other to put the accounting profession very much on the backfoot on some key debates.

Briloff himself was an important input into Criscione's biography, and in that sense it is an authorized one. However, many others were interviewed, and Criscione showcases the arguments of Briloff's critics as well his supporters. Chapter 7 covers Briloff's appearance before the AICPA Ethics Division Trial Board. To put it mildly, Briloff's personality and his umbrage at having his own integrity questioned did not assist him in these hearings. I found one of the most interesting aspects of the book to be Briloff's own response to a question by Criscione, re "how he now felt about the investigation and subsequent trial." Briloff answered that "he still felt the charges were absurd but he sympathized with the trial board's position. He mused, 'Would I do it all again? Would I go looking for an argument? No'" (p. 205).

Criscione's book provides much insight into an important figure who, over the 1965 to 2000 period (and beyond), has been a very consistent proponent of his view of the important dimensions of independence in the accounting profession. Changes made by regulators and more begrudgingly by the accounting profession itself have moved the industry structure sizably toward Briloff's position. A read of this book provides fascinating insights into an individual who both "pushed hard" and was "pushed hard against" as the accounting profession underwent some major (and to many within the profession) wrenching changes. It is well worth reading and will make an important contribution to our knowledge of the growth and development of the accounting profession in the last 40 years.

REFERENCES

Foster, G. 1979. Briloff and the capital market. *Journal of Accounting Research* 17 (1): 262–274.

——— 1987. Rambo IX: Briloff and the capital market. *Journal of Accounting, Auditing & Finance* 2 (4): 409–430.

GEORGE FOSTER
Professor of Management
Stanford University

WERNER DE BONDT (editor), *Financial Accounting and Investment Management* (Cheltenham, Glos, U.K.: Edward Elgar, 2009, ISBN: 978-1-84844-039-5, pp. xl, 609 (vol. I), viii, 526 (vol. II)).

In this two-volume set the reader is provided with an interesting collection of articles that relate to seven topics:

- Accountants and analysts as financial intermediaries (Volume I, Part I)
- Measuring corporate earnings and profitability (Volume I, Part II)
- Financial ratios, the risk of failure, and stock returns (Volume I, Part III)
- Equity valuation (Volume I, Part IV)
- Price-earnings ratios, market-to-book ratios, and stock returns (Volume II, Part I)
- Earnings and stock returns (Volume II, Part II)
- Fundamental analysis and stock returns (Volume II, Part III)

In addition, Professor Werner De Bondt, the editor, writes a set of insightful introductory comments in which he briefly discusses each topic, motivates and describes the articles he chose, and explains how the seven topics are linked together. This is especially helpful as it allows the reader to see the articles, many of which are well known, from Professor De Bondt's perspective. In particular, an important theme running through the introduction is that extant research provides considerable evidence that accounting numbers are related to equity prices and equity returns, and that they play a role in the risk-assessment process. However, the evidence also implies that, in certain circumstances, capital market participants do not fully understand the implications of past and current accounting numbers for expected future payoffs and outcomes. Hence, many unanswered questions remain regarding the investment-management role of financial accounting numbers.

The collection has many fine qualities, two of which stand out. First, when read together, the articles that relate to a particular topic nicely illustrate how our understanding of that topic has evolved, and they help the reader identify key research questions that remain unanswered. For example, in Volume I, Part I, which relates to the financial intermediation role of accounting and analysts' reports, Professor De Bondt begins with a 1964 article by the accountant Howard C. Greer, who describes many of the complaints that individual investors had (and continue to have) with the use and interpretation of accounting numbers (Greer 1964). Next, Professor De Bondt includes the seminal 1968 article by Ball and Brown, which demonstrates that, notwithstanding the complaints of individual investors, accounting earnings do have information content (Ball and Brown 1968). An immediate implication of this result is that there is a demand for earnings forecasts and, consequently, the analysts who make them. This clearly motivates the third article chosen by Professor De Bondt, which is by Asquith et al. (2005). These authors show that both the textual content as well as numerical earnings forecasts and price recommendations contained in a sample of *Institutional Investor* All-American analyst reports have information content. Finally, Professor De Bondt finishes Volume I, Part I with an article by de Franco et al. (2007), who show that some analysts provide biased information and that small investors do not always fully adjust for the bias. Hence, the reader finishes this part of the collection realizing that analysts' reports are informative on average but that we do not fully understand the factors that determine their quality or the manner in which capital-market participants use them when making investment decisions.

Second, even though each of the seven topics covered represent well-developed sub-fields of capital markets research, Professor De Bondt manages to link them together in a coherent fashion. He does this in two ways. First, he provides a concise yet illuminating description of the links between the topics in his introduction. Second, he chooses a set of articles that are very complementary. For example, each of the articles in Volume II, Part I, which relates to different valuation multiples (e.g., the price-to-earnings ratio) are clearly related to each of the six remaining topics. This is useful as it helps readers understand better the relation between various sub-fields and to draw connections between different bodies of evidence.

The collection also has some shortcomings, which mostly relate to the choice of articles selected and, thus, are unavoidable, given that there is a vast literature on the seven topics covered. Two of these shortcomings bear mentioning. First, all of the recent analytical work related to residual income valuation (e.g., Ohlson [1995]) and abnormal earnings growth valuation (e.g., Ohlson and Juettner-Nauroth [2005]) is omitted. Given the impact of this work on research and practice, the failure to include at least one of the key articles in this area is a limitation. Second, none of the recent articles regarding accounting-based estimates of the expected rate of return on equity capital are included. These estimates are becoming increasingly popular; hence, some discussion of why they are used and their properties seems warranted.

The collection will have value to a number of audiences. In my opinion, the individuals who will benefit the most are those with exposure to the fundamentals of finance and accounting (i.e., dividend policy and capital structure irrelevance, market efficiency, asset pricing models, etc.) but are not "black belts" in accounting research. Hence, Ph.D. students in accounting and finance who have progressed beyond their first year; faculty in accounting or finance who want a new perspective, and practitioners with a strong accounting and finance background will likely appreciate the two-volume set the most. However, because Professor De Bondt wisely chooses not to try to cover all seven topics from the ground up, individuals without some exposure to basic accounting and finance may find the two-volume set less useful.

REFERENCES

Asquith, P., M. B. Mikhail, and A. S. Au. 2005. Information content of equity analyst reports. *Journal of Financial Economics* 75 (2): 245–282.

Ball, R., and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (2): 159–178.

de Franco, G., H. Lu, and F. P. Vasvari. 2007. Wealth transfer effects of analysts' misleading behaviour. *Journal of Accounting Research* 45 (1): 71–110.

Greer, H. C. 1964. The corporate stockholder–accounting's forgotten man. *The Accounting Review* 39 (1): 22–31.

Ohlson, J. A. 1995. Earnings, book values and dividends in equity valuation. *Contemporary Accounting Research* 11 (2): 661–687.

———, and B. E. Juettner-Nauroth. 2005. Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies* 10 (2–3): 349–365.

STEVEN J. MONAHAN
Assistant Professor of Accounting
INSEAD

KEN McPHAIL and DIANE WALTERS, *Accounting and Business Ethics: An Introduction* (London, U.K.: Routledge, 2009, ISBN: 13:978-0-415-36235-1, pp. xii, 225).

This book is divided into two main parts, and divided further into nine chapters. Part I is "How to think ethically about accounting." It contains an introductory chapter, Chapter 2 on the psychology of moral judgment; Chapter 3 presenting a variety of ethical theories; Chapter 4 on the ethical foundations of accounting in terms of political philosophy; and Chapter 5 on the post- and new-modern approaches to accounting ethics. Part II is "The ethics of accounting practice." It consists of an analysis (Chapter 6) of the function of accounting in the context of "free" markets and of the morality of markets; Chapter 7 on accounting as a professional occupation; Chapter 8 on the ethics of international accounting; and a concluding chapter on ethics, intellectual capital, and accounting reporting. In addition to the text itself, each chapter contains a list of references and an extensive list of additional resources, including books, articles, podcasts, blogs, websites, and movies. A rough estimate is that one-third of the book is devoted to these reference materials. Since the actual text of the book is quite brief, roughly 140 pages (including a large number of illustrative boxes), the book is a short introduction to the issues and relevant literature, and is best viewed as a reference work rather than as a textbook *per se*.

Ken McPhail is professor of social and ethical accounting at the University of Glasgow, and Diane Walters is a lecturer at Heriot-Watt University, in Edinburgh.

This book is intended as a textbook for a course on accounting ethics. As the chapter topics indicate, the primary purpose of the book is to introduce accounting students to a wide range of issues and theories relating to the ethical status of the accounting profession. The discussion is at a "macro" level, for the most part; thus, it addresses issues relating to codes of ethics (e.g., why we have them, where they came from, and the extent to which an effective international code of ethics might be feasible), but does not go into detailed discussions of the contents of any specific code or of any specific ethical issue (such as auditor independence or conflict of interest). Possible uses of this book would be as a basis for a course in accounting ethics or as a supplement to such a course. In either case, the end-of-chapter lists of resources should provide much of this.

This review will proceed from the assumption that most accounting instructors in many countries will be relatively unfamiliar with most of the issues and theories covered in this book, but wish to learn about them for teaching or scholarly purposes. This suggests that the book may be useful to academics who wish either to teach some of the issues contained in this book, or to learn about them for their own scholarly development.

From the point of view of the mainstream of academic accounting thought (which in much of the world is dominated by economics, financial economics, and psychology), the potential value of this book is to present a number of ethical issues and theories that have been given little attention in North America and other places. Thus, although a number of more familiar theories are presented in Chapters 2, 3, and 4, the heart of Part I is the presentation of the views of a variety of lesser-known (from the standpoint of mainstream academics) social theorists, including Jean-Jacques Rousseau, Jürgen Habermas, Emmanuel Levinas, Jacques Derrida, and Michel Foucault. Following this, Part II presents a number of fundamental issues that, according to the authors, face the accounting profession, but which have been ignored by the mainstream. Chapter 2 on moral psychology seems out of place, in part because it focuses on the ability of accountants (and accounting students) to deal with day-to-day ethical issues, and does not play an important role in the rest of the book.

Unfortunately, the brevity of the text limits the book's usefulness as a textbook for most instructors, at least in the North American context, because students will have given little or no thought to either the theories or the issues discussed in the book. It also limits its usefulness for accounting academics, many of whom will also be neophytes. The problem is repeated many times in the book. For example, Habermas' theory is described in three pages (Chapter 5), Foucault's theory in two pages (Chapter 5), a little over three pages are devoted to Levinas (Chapter 4), Kant is covered in one page (Chapter 3), and only a half-page treats Rawls (Chapter 3). (Less space is given to Rawls than to a newspaper article by Prem Sikka [Chapter 4].) An important pair of ethical concepts generally, and for accounting in particular, are rights and duties. Rights (but not duties) are dealt with at somewhat greater length (Chapter 4), but even here the agenda of Part II focuses attention on human rights, rather than on, or in addition to, rights (e.g., rights to privacy or property) that are relevant to the issues practicing accountants face in their day-to-day work.

The fundamental problem here is that neophytes wanting an introduction to this literature are likely to gain little from these discussions of complex and subtle theories. Academics who are somewhat familiar with the work of Anglo-Saxon (e.g., Locke, Hobbes, Kant, Mill, and Rawls) or of Continental philosophers (e.g., Levinas, Derrida, and Foucault) may be able to fill in the blanks themselves by searching out other sources, but will not gain much here. Without prior understanding, the accounts given here will be of little help. For example, the authors quote from a secondary source about the essence of Levinas' view: "Morality begins in the face to face," and, "morality is the encounter with the Other as face" (p. 75). "What on earth does that mean?" the authors ask; unfortunately, the explanation does not clarify it in a way that an "outsider" to this literature will understand.

More specifically, the text in Chapter 3, devoted to the major ethical theories, is 11 pages long, including about four pages of boxes on such topics as Al Gore, Dirty Harry, circle time, thalidomide, and Islamic accounting. Even having radically trimmed the types and variations of theories discussed, it is impossible to say much that is intelligible in such a brief space. And even then, the half-page devoted to Rawls is not only short but fundamentally mistaken in suggesting that a practicing accountant should deal with a material error by employing the original position argument. For one thing, the authors do not mention Rawls's (1971) famous principles of justice. Even though they refer to Rawls's famous 1955 paper on rules (which is worth reading) in another connection, they do not recognize the importance of rules, which would be an important feature of his approach to the example of an erroneous financial statement (Rawls 1955). (The authors do discuss "rules-based" versus "principle-based" standards, but not in this context.)

In addition, these incomplete discussions are often left for the reader to fill in. For example, the discussion of "knowledge capitalism" in Chapter 9 suggests that it provides the possibility for "empowerment and a strengthening of the 'democratic public sphere'" (p. 185). The authors then suggest that the reader "[t]ry to link this kind of perspective back to our discussion of Jürgen Habermas in Chapter 5. Think about how and whether new technological forms contribute towards the kind of communicative processes that Habermas promotes." This would be very difficult to do, based only on the three-page discussion of Habermas in Chapter 5. Rather than leaving this to the reader to figure out, the book should have explained this linkage, with the attendant possibility of increasing readers' understanding of Habermas' theory.

The theories presented in Part I may be viewed as the foundations of Part II, which directly addresses a number of fundamental ethical issues relating to the practice of accounting. For this reason, readers may find Part II more interesting, from the point of view of understanding the ethical foundations of accounting.

Chapter 6 discusses the supposed ethical neutrality of accounting. The basic idea is that accounting today, for the most part, occurs in the context of market-based capitalist economic systems, and that the ethical justifiability of accounting presumes the ethical status of such economic systems. For example, accounting is practiced in the context of a specific configuration of rights (property rights) that give priority to capital over labor, and so the defensibility of this type of accounting depends on the defensibility of these rights. In this sense, accountants

would need to take responsibility for their role in such an economic system. This is unexceptionable. Because accounting (whatever else it is) is supposed to be a practical occupation, it is successful to the extent that it serves the purposes set out for it. The important observation is that accounting is not "value-neutral." The authors' main concern is an important one: to consider the ethical status of the activity and institution to which accounting contributes.

The main problem with this chapter is that it finds the ethical foundations of capitalism in a form of utilitarianism. This is surprising, because it seems clear that many ethical theories (including utilitarianism) can provide the ethical underpinnings of capitalism. An obvious example is libertarianism. Indeed, a major problem with this chapter is that, while it recognizes that the mainstream of accounting thought is capitalist in some sense, it does not engage the advocates of this view, who would agree about the centrality of market economies, and would emphasize (and defend) the use of accounting in securities markets.

Chapter 7 examines accounting as a profession, focusing on the place of the profession in society, rather than on specific aspects of the practice of accounting. Of special interest is the important question whether accounting, as a professionalized occupation, acts in the public interest (as is often claimed) or in the interest of private groups. (Whether accounting is a profession at all is addressed only briefly.) For example, do codes of ethical conduct serve the interests of the profession or of society? As with Chapter 6, it would have been good for the authors to have discussed the more mainstream literature on these topics, because many "mainstream" academics are just as suspicious as the authors appear to be of the public-interest claims of the profession. It would have been interesting to see just what the differences are between the authors' and the mainstream points of view.

Chapter 8 addresses international issues in the practice of accounting. The authors correctly point out that the desirability and feasibility of trying to enforce uniform standards of practice (whether GAAP, GAAS, or ethical standards) present major difficulties. Issues include respect for cultural tradition and the possibility that a particular set of economic and ethical practices may be imposed on cultures that have developed different traditions. As in Chapter 6, it is observed that international harmonization is not ethically neutral, and that accountants implicitly act in the interest of whatever economic system or ideology is dominant. This is fairly obvious, and indeed many accountants and academics would defend the current global system and the role of accountants in it. The contribution of the authors is to claim that this is contestable and not as obvious as some might think.

The authors appear to take a strong view on this issue, criticizing both the impact of globalization on less developed economies (the activities of the World Bank attract special attention) and the activities of accountants in helping to advance a globalization agenda. In particular, they claim that accounting may have a role in the increase in international terrorism because of the subjugation of "indigenous" people. For example, they repeat an argument that accountants "broadly defined" in Canada have been responsible for First Nations children being sent to boarding schools, for the murder of First Nations people in order to take their land (because accountants paid the bounty to the killers), and the giving of gifts to First Nations people. Although the idea of role morality (i.e., the morality of people acting in their role as accountant) is discussed in passing elsewhere, the analysis here does not recognize that the ethical status of taking children from their parents or killing people for their land may be wrong, and that an accountant who takes part in this is ethically culpable (as a person); but this does not imply that the wrongness of the act is due to the precise occupational role (i.e., the accountant as accountant) that a person plays.

Chapter 9, like Chapter 2, is not a good fit with the rest of the book, as I interpret the authors' mission. It is not clear what intellectual capital, and accounting for it, has to do with the ethics of accounting. However, the authors consider whether "ethical knowledge" should be considered an asset of a company. They describe one research study that suggests that behaving ethically is more profitable; but they ignore a very large literature, the results of which are mixed. In any case, many ethicists question the legitimacy of such a linkage, because emphasizing it suggests that the reason to be ethical is to produce profits.

In conclusion, this book has some attractive features. First, it may serve as an introduction to a number of fundamental ethical issues of the accounting profession. This may be helpful to academics, in relation either to teaching students or to advance their own scholarly understanding of the profession. However, the text is brief and sketchy, sometimes to the point of being uninformative (and mistaken). This means that a neophyte will get, at best, a very rough outline of the issues and the literature, and someone who is familiar with the literature will not obtain additional perspective or understanding. In addition, the early chapters on ethics and social and political philosophy leave out a number of important theories that are relevant to the ethics of the accounting profession.

The extensive reference materials at the ends of the chapters are a strength of the book, in large part because much of the material presented in this book will be unfamiliar to most academics in many countries. The teaching-oriented resources (blogs, podcasts, films, etc., as well as the books and articles) may be particularly valuable. Adoption of this book for a course would require heavy supplementation, possibly with the assignment of some of these resources. However, the resources are incomplete, because they are a function of the choices made by the authors about what to include in the book. In view of the many omissions, this is an important limitation of its value as a reference work.

A few quibbles: (1) the authors sometimes write in the first person. This is OK, except that sometimes it is singular and sometimes plural. And (2) the book contains a large number of text boxes (many of them up to a page long) that discuss various issues or provide background information. Unfortunately, many of them add little or nothing to gaining an understanding of the issue or theory in question. In fact, I found many of them to be unhelpful and even distracting. Indeed, it is unclear what the point is of many of the boxes. In addition, the material contained in the boxes is almost always not given a reference source. While it may be possible to figure out from the end-of-chapter references where the information came from, it should not be necessary. Less of a quibble is that their number and length limit the number of pages available for discursive text.

REFERENCES

Rawls, J. 1955. Two concepts of rules. *The Philosophical Review* 64 (1): 3–32.
——— 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

JAMES C. GAA
Professor of Accounting
University of Alberta

OTHMAR W. WINKLER, *Interpreting Economic and Social Data: A Foundation of Descriptive Statistics* (Berlin, Germany: Springer, 2009, ISBN 978-3-540-68721-4, pp. xvi, 265).

Opening this book by Othmar Winkler is like splashing oneself with cold water at 5:30 in the morning. It's a wakeup call! The author lays out his "call to arms" in the preface. In our quest to understand or "make sense of socio-economic data" (p. vi), we have come to rely too heavily on statistical inference (F, t, Chi-square) and on assumed symmetry and continuity. If we seek insights, we are enjoined to adopt instead the descriptive tools of statistics and apply them to aggregate observations and their categorizations. To understand socio-economic phenomena, it is essential to recognize that the contributing processes are purposeful and not random.

Winkler appropriately observes that inferential statistical methods rose to prominence in the natural sciences. But these methods are typically ill-suited when adopted to grapple with socio-economic data. He takes pains to point out the limitations when directly transporting inferential statistical methodology to economic and social data. Among the more readily identified differences is the fact that social science data occur over time and location and are subject-specific. Insight into the data generation process commands that the observer consider the data's uniqueness to each of these factors.

Games of chance are offered as an ideal setting for the probability calculus. The rules of the game are known and cannot change during the game. The trials are repeated identically, and outcomes can be identified and predicted over the long run. Economic games, in business for example, are very different animals. Essentially, none can be accurately described as being repeated under stable conditions with known outcomes and subject to long-term predictability. Over time, conditions change, cooperation may develop, and new options typically emerge. Outcomes are multidimensional. Perturbations and changes over time present serious difficulties for aggregation into a single measure of success. Indeed, success is likely to be a vector, not a datum. Winkler writes, "It simply goes too far to view economic processes as random experiments" (p. 3). One need only cast an eye as far as the tightly controlled economies of the former European Soviet-bloc countries, says Winkler (p. 187), to observe the lack of randomness. Could we deny government influence even in market-based economies? He observes, "Rather [chance] resides in the observer's difficulty, inability or lack of interest to understand the deterministic causal system that actually is at work" (p. 189).

Of equal if not greater importance is the fact that randomness plays a very different role across disciplines. In games of chance and in the sciences, randomness is captured by the probability calculus. It underlies experimental design. In the socio-economic world, the remaining "random" observations after appropriate aggregation and classification often are what are of most interest, because the objective is to understand the process generating the observations. In addition, and again as the author points out, there are "[v]ery few statistical observations ... made directly by an objective outside observer" (p. 5) essential to the randomness assumption.

Accountants will discover in Chapter 11 that many of the statistical tools favored by the author have been and continue to be used in accounting. He points out that the collection, classification, and aggregation (and de-aggregation) of data that are involved in preparing the balance sheet and income statement are simply standard statistical estimation processes. The data can then be analyzed horizontally or vertically, as is typically taught to students. The consideration of the time-series nature of the properly categorized measurements leaps out at the reader. But the complexities of the traditional statistical analysis of trend, seasonal, and randomness are avoided. Inventories, sales, and income all yield to simple but different descriptive approaches. The observed extensive reliance on ratios in both accounting and financial analysis is a natural application of the descriptive statistical techniques extolled by the author. In addition, methods for improved ratio analysis are noted. The interested reader is left with the discovery of additional connections as well as a brief digression into auditing. With the exception of budgeting and record keeping, management accounting is ignored, a limitation for accountants but of minor importance to the text.

The author exhorts his readers to use descriptive statistical tools to understand, evaluate, and interpret social and economic activity (e.g., business formation, entrepreneurship, pricing, trade, production, poverty, and so on). He would have one's tool kit packed with the proper use of ratios, appropriate graphing and tabling presentation techniques, and frequency tabulations (including their use to disclose easily understood measures of dispersion). He urges us to aggregate and classify data as a means of understanding socio-economic phenomena, thereby allowing for the less tractable and measurable aspects of reality to be captured and interpreted. And, if I may, we should all adopt a critical mindset when evaluating the application of inferential statistics to the socio-economic literature which we read or disseminate. I believe the author would approve of the insights that Professors Steven D. Levitt and Stephen J. Dubner bring to investigating interesting topics in the social sciences, which have been widely popularized in their books, *Freakonomics* (2005) and *SuperFreakonomics* (2009).

The book contains many examples typically gathered from a career of studying, researching, and consulting that describe the problems of forcing an analysis into the statistical inference mold. One example that appears in Chapter 9 and several other chapters, including the author's "Afterthoughts" (pp. 229–232), describes a study using regression analysis for a trial centered on male/female pay differentials. When the results appeared to be counter-intuitive, the explanation resided in the socio-economic nature of the data. The interpretation of the slope coefficient had to be interpreted in the light of the changes in the data over time rather than as simply the average change in X for a unit of change in Y. The chapter examines many of the limitations of regression analysis as typically applied to socio-economic data.

The author suggests that our reliance on the inferential aspects of mathematical statistics comes from its elegance and our own desire to be able to run with the elite. In our research, there seems to be a quest for mathematical sophistication, as exemplified by advanced modeling and mathematical testing to satisfy reviewers. We as instructors find statistical inference more exciting to teach and more likely to reap the admiration of our colleagues. The author also points to the academic hiring function as seeking those with expertise and interests devoted to statistical inference where the random sample is king.

The guilt lies heavily with those of us who have authored or teach from the standard statistical texts. Too often we have not only undervalued the direct relevance of descriptive methods, but also have failed to identify, let alone appreciate, the inherent limitations underlying the applicability of assuming randomness and the associated applicability of statistical inference.

This book does not fit easily into most schools' curriculum. However, this result is to a large extent a problem with those schools' curriculum rather than with the book's subject matter. It is a short volume consisting of 12 chapters and less than 150 pages of text. The text is augmented by enough text notes—often containing their own rewarding gems—to recall the *Principles of Economics* text written by Alfred Marshall (1890) that many of us were required to digest in our economics classes. (By the way, Marshall respected mathematics but did not want it to dominate economics.) Many examples, culled from Winkler's extensive writings and experience, are sprinkled through the text and the aforementioned notes. Most of these are from another era, which makes them no less relevant today. The writing is clear, if wordy. (The author feels repetition is useful for those who may wish to use portions of the volume without missing the essence of the argument.)

Indeed, statistics classes might be more exciting if students were given questions or issues to investigate, similar to the pay example explored by the author and described in some detail in Chapter 9. Discovering methods that yield insights and that reveal the limitations of investigative techniques might be rewarding and actually fun. Accounting professors: take note.

If your school or department has a seminar that dwells on the philosophy of science, treats research techniques in the social sciences, or just reviews research studies using statistics, then a reading of portions (or all) of this book would be both appropriate and desirable. If not, I would urge the Dean (or Department Chair) to provide each professor, instructor, and doctoral student, regardless of their field of study, a copy at the holidays and then hold an exam the day before classes begin.

REFERENCES

Levitt, S. D., and S. J. Dubner. 2005. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York, NY: HarperCollins.

———, and ———. 2009. *SuperFreakonomics: Global Cooling, Patriotic Prostitutes and Why Suicide Bombers Should Buy Life Insurance*. New York, NY: HarperCollins.

Marshall, A. 1890. *Principles of Economics*. London, U.K.: Macmillan.

THOMAS R. DYCKMAN
Emeritus Professor of Accounting
Cornell University

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature. Topical areas of interest to the journal include accounting information systems, auditing and assurance services, financial accounting, management accounting, taxation, and all other areas of accounting, broadly defined. The journal is also open to all rigorous research methods.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and editors are responsible for providing constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be formatted in 12-point font on 8 1/2 x 11″ paper and should be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins should be at least one inch from top, bottom, and sides.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we".
5. A cover page should show the title of the paper, all authors' names, titles and affiliations, email addresses, and any acknowledgments.
6. The American Accounting Association encourages use of gender-neutral language in its publications.

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a cross-sectional equation; re-form*. See *Webster's* for correct usage.

Keywords: The abstract must be followed by at least three keywords to assist in indexing the paper and identifying qualified reviewers.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words (150 maximum) should be presented on a separate page immediately preceeding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. When the reference list contains two or more works by the same author (as the only author or first of two or more authors) published in the same year, the suffix a, b, etc., is appended to the date in the within-text citations and in the "References" section. For example, (Johansson 2004a, 2004b, 2004c; Baiman and Rajan 2002a, 2002b; Dhaliwal et al. 2005a, 2005b).
4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works should use acronyms or short titles where practicable. For example: (NCFFR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in *alphabetical order* according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Use authors' initials instead of proper names.
3. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
4. Date of publication follows the author's (authors') name(s).
5. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

———, and ———. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.

Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: http://hsgac.senate.gov/files/111803berry.pdf.

Cohen, D., A. Dey, and T. Lys. 2005. The Sarbanes Oxley Act of 2002: Implications for compensation structure and risk-taking incentives of CEOs. Working paper, New York University, University of Chicago, and Northwestern University.

Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition.* Ann Arbor, MI: University of Michigan.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

———, and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.

Dhaliwal, D., M. Erickson, and O. Li. 2005a. Shareholder income taxes and the relation between earnings and returns. *Contemporary Accounting Research* 22: 587–616.

———, L. Krull, O. Li, and W. Moser. 2005b. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43: 675–708.

Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.

Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109.* FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.

Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes.* Boston, MA: Houghton Mifflin.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.

Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics* (Winter) 30: 555–574.

National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: NCFFR.

Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.

———. 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy*, edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.

Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.

Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports.* Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be inserted using the "footnote" or "endnote" feature of the word processing software which will automatically number the footnotes throughout the manuscript with superscript Arabic numerals.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state upon submission that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. Manuscripts should be submitted via email as Microsoft Word or PDF file to the Senior Editor, Steven Kachelmeier, at email address tar@mccombs.utexas.edu. Please submit separate files for (1) the manuscript's title page with identifying information (not forwarded to reviewers), (2) the manuscript with title page and all other identifying information removed, and (3) any necessary supplemental files, such as experimental instructions and/or response memoranda on invited revisions. The body or subject line of the email should also indicate the submission fee receipt code from AAA (see step 4 below).
4. A nonrefundable submission fee is required unless explicitly waived by the editor for invited revisions of previous submissions. The submission fee in U.S. funds is $200.00 for members and $400.00 for nonmembers of the AAA payable by credit card (VISA or MasterCard only). The payment form is available online at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If you are unable to pay by credit card or have any questions, please contact the AAA Member Services Team at (941) 921-7747 or info@aaahq.org.
5. Revisions must be submitted within 12 months from the decision letter inviting a revision.

COMMENTS

The Accounting Review welcomes submissions of comments on previous *TAR* articles. Comments on articles previously published in *The Accounting Review* will generally be reviewed by two reviewers, usually including an author of the original article (to assist the editor in evaluating whether the submitted comment represents the prior article accurately) and an additional reviewer who is independent of the original article. If a comment is accepted for publication, the original author will generally be invited to reply. All other editorial requirements, as enumerated above, apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The AAA's Executive Committee policy (originally adopted in 1989, and amended in 2009) is that the objective of the Association-wide journals (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. To fulfill this objective, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles.

THE ACCOUNTING REVIEW

VOLUME 85, No. 6 NOVEMBER 2010

JOHN CHRISTENSEN
Accounting Errors and Errors of Accounting

STANLEY BAIMAN, SERGUEI NETESSINE, and RICHARD SAOUMA
Informativeness, Incentive Compensation, and the Choice of Inventory Buffer

JASMIJN C. BOL, TIMOTHY M. KEUNE, ELLA MAE MATSUMURA, and JAE YONG SHIN
Supervisor Discretion in Target Setting: An Empirical Investigation

GAVIN CASSAR and JOSEPH GERAKOS
Determinants of Hedge Fund Internal Controls and Fees

MERLE EDERHOF
Discretion in Bonus Plans

MEI FENG and ADAM S. KOCH
Once Bitten, Twice Shy: The Relation between Outcomes of Earnings Guidance and Management Guidance Strategy

PAUL E. FISCHER and PHILLIP C. STOCKEN
Analyst Information Acquisition and Communication

KIRIDARAN KANAGARETNAM, GOPAL V. KRISHNAN, and GERALD J. LOBO
An Empirical Analysis of Auditor Independence in the Banking Industry

EDMUND C. KEUNG
Do Supplementary Sales Forecasts Increase the Credibility of Financial Analysts' Earnings Forecasts?

KRISHNAGOPAL MENON and DAVID D. WILLIAMS
Investor Reaction to Going Concern Audit Reports

BRIAN P. MILLER
The Effects of Reporting Complexity on Small and Large Investor Trading

ANDREW B. REFFETT
Can Identifying and Investigating Fraud Risks Increase Auditors' Liability?

Marygrove College Library
8425 West McNichols Road
Detroit, MI 48221

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2010–2011 Executive Committee of the Association are as follows:

President	Kevin D. Stocks, Brigham Young University
President-Elect	Gregory B. Waymire, Emory University
Past President	Nancy A. Bagranoff, Old Dominion University
Vice President	Ira Solomon, University of Illinois at Urbana–Champaign
Vice President–Education	Belverd E. Needles, Jr., DePaul University
Vice President–Education-Elect	Alan Reinstein, Wayne State University
Vice President–Finance	Stacy E. Kovar, Kansas State University
Vice President–International	Kazuo Hiramatsu, Kwansei Gakuin University
Vice President–Professional Relations	Robert H. Colson, Grant Thornton, LLP
Vice President–Publications	Jean C. Bedard, Bentley University
Vice President–Publications-Elect	Christine A. Botosan, The University of Utah
Vice President–Research	John Christensen, University of Southern Denmark
Vice President–Sections and Regions	Gail Hoover King, Rockhurst University

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues include *Accounting Education News*, and online selection of one, two, or all three of the following Association-wide journals: *The Accounting Review*, *Accounting Horizons*, and *Issues in Accounting Education*. (This does not include Section journals.) Select from:

The Accounting Review only	U.S. $185.00	with Print $220.00
Accounting Horizons OR *Issues in Accounting Education* only	U.S. $185.00	with Print $200.00
The Accounting Review AND *Accounting Horizons* OR *Issues in Accounting Education*	U.S. $195.00	with Print $245.00
Accounting Horizons AND *Issues in Accounting Education*	U.S. $195.00	with Print $225.00
All three Association-wide journals	U.S. $205.00	with Print $270.00

Associate membership:
Full-time students are eligible for associate membership. The print option is not available for Associate members. Select from:

Any one Association-wide journal	U.S. $25.00
Any two Association-wide journals	U.S. $35.00
All three Association-wide journals	U.S. $45.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Auditing; Diversity; Financial Accounting and Reporting; Forensic and Investigative Accounting; Gender Issues and Worklife Balance; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Strategic and Emerging Technologies; Teaching, Learning and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to tar@mccombs.utexas.edu. If it is necessary to send any hard copy correspondence, the mailing address is Mary Capps, Editorial Assistant, *The Accounting Review*, The University of Texas at Austin; Department of Accounting, CBA 4M202; 1 University Station B6400; Austin, TX 78712-0211 • Telephone: (512) 471-5318 • Fax: (512) 471-3904.

The submission fee of $200.00 for members of the AAA or $400.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Info@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

The Accounting Review is indexed in the Social Science Citation Index which is included in Web of Science.

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rates of $375 for a print copy or at $425 for a print and electronic copy. The member subscription rate for a hard copy is $35. Electronic access is included with membership dues. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review*, 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2010 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Steven J. Kachelmeier

Editors

Shannon W. Anderson
J. Harry Evans III
James E. Hunton
Kathryn Kadous
Sanjay G. Kallapur
Ranjani Krishnan
Laureen A. Maines
D. Paul Newman
Thomas C. Omer
Wayne B. Thomas
Mark A. Trombley
Beverly R. Walther
Paul A. Zarowin

VOL. 85 NOVEMBER 2010 No. 6

JOHN CHRISTENSEN
Accounting Errors and Errors of Accounting **1827**
STANLEY BAIMAN, SERGUEI NETESSINE, and RICHARD SAOUMA
Informativeness, Incentive Compensation, and the Choice of Inventory Buffer ... **1839**
JASMIJN C. BOL, TIMOTHY M. KEUNE, ELLA MAE MATSUMURA, and JAE YONG SHIN
Supervisor Discretion in Target Setting: An Empirical Investigation **1861**
GAVIN CASSAR and JOSEPH GERAKOS
Determinants of Hedge Fund Internal Controls and Fees **1887**
MERLE EDERHOF
Discretion in Bonus Plans **1921**
MEI FENG and ADAM S. KOCH
Once Bitten, Twice Shy: The Relation between Outcomes of Earnings Guidance and Management Guidance Strategy **1951**
PAUL E. FISCHER and PHILLIP C. STOCKEN
Analyst Information Acquisition and Communication **1985**
KIRIDARAN KANAGARETNAM, GOPAL V. KRISHNAN, and GERALD J. LOBO
An Empirical Analysis of Auditor Independence in the Banking Industry **2011**
EDMUND C. KEUNG
Do Supplementary Sales Forecasts Increase the Credibility of Financial Analysts' Earnings Forecasts? **2047**
KRISHNAGOPAL MENON and DAVID D. WILLIAMS
Investor Reaction to Going Concern Audit Reports **2075**
BRIAN P. MILLER
The Effects of Reporting Complexity on Small and Large Investor Trading **2107**
ANDREW B. REFFETT
Can Identifying and Investigating Fraud Risks Increase Auditors' Liability? ... **2145**

BOOK REVIEWS, Stephen A. Zeff, Editor

Developments in New Reporting Models ..... CHRISTINE A. BOTOSAN **2169**

Annual Report and Editorial Commentary for *The Accounting Review* **2173**
Editorial Policy and Style Information **2205**
Index **2209**

Editorial Staff of *The Accounting Review*
2010–2011

SENIOR EDITOR
Steven J. Kachelmeier, The University of Texas at Austin

EDITORS
Shannon W. Anderson, University of California, Davis
J. Harry Evans III, University of Pittsburgh
James E. Hunton, Bentley University
Kathryn Kadous, Emory University
Sanjay G. Kallapur, Indian School of Business
Ranjani Krishnan, Michigan State University
Laureen A. Maines, Indiana University
D. Paul Newman, The University of Texas at Austin
Thomas C. Omer, Texas A&M University
Wayne B. Thomas, University of Oklahoma
Mark A. Trombley, The University of Arizona
Beverly R. Walther, Northwestern University
Paul A. Zarowin, New York University

EDITORIAL ASSISTANT
Mary C. Capps, The University of Texas at Austin

RESEARCH ASSISTANT
Tracie Majors, The University of Texas at Austin

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland

COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Lisa Habblitz
Lindsay Wharton
Kathy Casper
Judy Cothern
Sonja Langley
Jan Kovarik (subcontracted)
Copyediting Team (subcontracted)

OTHER AAA STAFF
Dierdre Harris
David Boynton
Alex Metz
Pat Stein
Debora Gardner
Peggy Turczyn
Barbara Gutierrez
Barbara Brady
Arlene Oliver
Roland LaTulip
Diane Leger
Man Fong
Jodie M. Morrill
Michelle Bowles
Kurt Gardner

Editorial Advisory and Review Board

Anwer S. Ahmed, Texas A&M University
Ashiq Ali, The University of Texas at Dallas
Anil Arya, The Ohio State University
Benjamin C. Ayers, University of Georgia
William R. Baber, Georgetown University
Stephen P. Baginski, University of Georgia
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Linda Smith Bamber, University of Georgia
Orie E. Barron, The Pennsylvania State University
Anne L. Beatty, The Ohio State University
Jean C. Bedard, Bentley University
Robert J. Bloomfield, Cornell University
Jennifer L. Blouin, University of Pennsylvania
Mark T. Bradshaw, Boston College
Lawrence D. Brown, Georgia State University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
Joseph V. Carcello, The University of Tennessee
Mary Ellen Carter, Boston College
Paul K. Chaney, Vanderbilt University
Qiang Cheng, University of Wisconsin–Madison
Michael B. Clement, The University of Texas at Austin
William M. Cready, The University of Texas at Dallas
Antonio Davila, IESE Business School–University of Navarra
Patricia M. Dechow, University of California, Berkeley
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Sunil Dutta, University of California, Berkeley
Leslie G. Eldenburg, The University of Arizona
Patricia M. Fairfield, Georgetown University
David B. Farber, University of Missouri–Columbia
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
James R. Frederickson, Melbourne Business School
Frank B. Gigler, University of Minnesota
Cristi A. Gleason, The University of Iowa
Johnathan C. Glover, Carnegie Mellon University
Steven M. Glover, Brigham Young University
David A. Guenther, University of Oregon
Luzi Hail, University of Pennsylvania
Jeffrey Hales, Georgia Institute of Technology
John R. M. Hand, The University of North Carolina at Chapel Hill
R. Lynn Hannan, Georgia State University
Gilles Hilary, INSEAD
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
Chris E. Hogan, Michigan State University
Ole-Kristian Hope, University of Toronto
Richard W. Houston, The University of Alabama
Raffi J. Indjejikian, University of Michigan
Christopher D. Ittner, University of Pennsylvania
Debra C. Jeter, Vanderbilt University
Karla M. Johnstone, University of Wisconsin–Madison
Bin Ke, Nanyang Technological University
S. Jane Kennedy, University of Washington
Inder K. Khurana, University of Missouri–Columbia
Michael D. Kimbrough, University of Maryland

Ronald R. King, Washington University in St. Louis
April Klein, New York University
W. Robert Knechel, University of Florida
Lisa L. Koonce, The University of Texas at Austin
Susan D. Krische, American University
Jayanthi Krishnan, Temple University
William N. Lanen, University of Michigan
David F. Larcker, Stanford University
Reuven Lehavy, University of Michigan
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Joan L. Luft, Michigan State University
Anne M. Magro, George Mason University
Carol Marquardt, Baruch College–CUNY
Dawn A. Matsumoto, University of Washington
Edward Maydew, The University of North Carolina at Chapel Hill
Brian W. Mayhew, University of Wisconsin–Madison
Sarah E. McVay, University of Utah
Kenneth A. Merchant, University of Southern California
Gregory S. Miller, University of Michigan
Lillian F. Mills, The University of Texas at Austin
Brian Mittendorf, The Ohio State University
Partha S. Mohanram, University of Toronto
Steven J. Monahan, INSEAD
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Arkansas
Venky Nagar, University of Michigan
Terry L. Neal, The University of Tennessee
Karen K. Nelson, Rice University
Edward F. O'Donnell, Southern Illinois University Carbondale
Per M. Olsson, Duke University
Chul W. Park, The University of Hong Kong
Mark E. Peecher, University of Illinois at Urbana–Champaign
Sean A. Peffer, University of Kentucky
Kathy Ruby Petroni, Michigan State University
John D. Phillips, University of Connecticut
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, Stanford University
Marlene A. Plumlee, University of Utah
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shiva Rajgopal, Emory University
Scott Richardson, BlackRock Inc.
John R. Robinson, The University of Texas at Austin
Darren T. Roulstone, The Ohio State University
Katherine Schipper, Duke University
Karen L. Sedatole, Michigan State University
Catherine M. Shakespeare, University of Michigan
Lakshmanan Shivakumar, London Business School
K. (Shiva) Sivaramakrishnan, University of Houston
Hollis A. Skaife, University of Wisconsin–Madison
Richard G. Sloan, University of California, Berkeley
J. Reed Smith, Indiana University at Indianapolis
Mark T. Soliman, University of Washington
Geoffrey B. Sprinkle, Indiana University
Mary Harris Stanford, Texas Christian University
Phillip C. Stocken, Dartmouth College
K. R. Subramanyam, University of Southern California
Hun-Tong Tan, Nanyang Technological University
Kristy L. Towry, Emory University
Senyo Tse, Texas A&M University
Jennifer Wu Tucker, University of Florida
Irem Tuna, London Business School
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Michael S. Wilkins, Texas A&M University
Richard H. Willis, Vanderbilt University
Arnold M. Wright, Northeastern University
Teri Lombardi Yohn, Indiana University
Xiao-Jun Zhang, University of California, Berkeley
Mark F. Zimbelman, Brigham Young University

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1959–1962	Robert K. Mautz
1962–1965	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William Kinney
1990–1994	A. Rashad Abdel-khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber
2002–2005	Terry Shevlin
2005–2008	Dan S. Dhaliwal

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 1827–1838

American Accounting Association
DOI: 10.2308/accr.2010.85.6.1827

Accounting Errors and Errors of Accounting

John Christensen
University of Southern Denmark

ABSTRACT: Accounting should pay more attention to errors, as errors are essential for the updating of beliefs. Accounting is an information system, and errors are the carriers of information according to Bayes' Theorem. Accountants are primarily concerned with the mean (value), but the variance of accounting numbers is equally important. Only autocorrelation makes historical accounting relevant for decision purposes. Endogenous errors of accounting are more common than acknowledged. First, the accounting model is linear, whereas the world is nonlinear. Second, accounting is not the only information channel, and accountants must consider the role of accounting when it supplements other information sources. This commentary discusses the consequences of endogenous errors. Errors are inherent to accounting, and accountants must address them.

Keywords: *information perspective; accounting errors; permanent earnings; nonlinearity; alternative information sources.*

I. INTRODUCTION

The accounting system is a closely managed information system. Financial data about the firm are collected with the purpose of guiding decisions related to the firm. The recognition rules govern the inclusion of information in the accounting library. Recognition rules also govern the exclusion of data from the accounting system. The accounting system is based upon old technology dating back to Pacioli in 1494. This technology continues to be used, and the market has proved it useful. Basically, it is a double entry system that simultaneously keeps track of the financial stocks and flows of the company, with the comparative advantage that it is hard to manipulate. Accountants have a long track record in maintaining the accounting system based upon financial data.

Even with this inherent security, the accounting system is by no means flawless (cf. Beaver 1998). The net result is that not all information of relevance for economic decisions is included in the accounting system. The accounting system is effectively limited to information that can be measured reliably. Thus, all troublesome items are carefully excluded, and, as a consequence, the accounting system steps back from the valuation of many intangible assets and writes off such investments at acquisition.

Helpful comments from Joel Demski, Steven Kachelmeier, and seminar participants at Bocconi University are gratefully acknowledged.

Editor's note: This commentary, based on a lecture at the 2009 American Accounting Association Annual Meeting in New York, NY, was invited by senior editor Steven Kachelmeier, consistent with the AAA Executive Committee's goal to promote broad dissemination of the AAA Presidential Scholar Lecture.

Submitted: May 2010
Accepted: June 2010
Published Online: November 2010

The accounting system serves many purposes in a multi-person world. A conflict among uses is inevitable, as pointed out by Gjesdal (1981). Thus, errors are inevitable. More to the point is that it is an information system, and information is imbedded in the errors. Hence, the question of how to extract information from the endogenous errors of the accounting system must be addressed. This brings the time-series properties of accounting numbers into focus. Extracting permanent and transitory earnings components from errors becomes an important exercise, as I discuss in Section II.

Costing and the managerial accounting system also contain endogenous errors. Most of these errors originate from the discrepancy between the accounting model and the underlying economics of the firm. The old variable-cost-versus-fixed-cost controversy might be seen as belonging to this type of error. The bigger picture is that the accounting model is linear, whereas the world is hardly linear. Economies of scale and scope are part of most businesses, and they certainly form the basis for most multi-product firms. The discrepancy between the linear accounting model and the nonlinear economics induces errors into the accounting system and reduces its usefulness for decision purposes. The activity-based costing (ABC) accounting system offers a finely tuned system, which mirrors the production technology of the firm, albeit using a linear costing technique. The alternative is a simple accounting system using fewer cost drivers. Both are infected by errors, although it is not clear that one should favor the more complex system. I return to this issue in Section III.

Another set of errors of the accounting system relates to the fact that accounting is constructed as if it were the only information source providing information to decision makers. Thus, much effort is directed to ensuring that the accounting system reflects all available information. For example, consider the fair-value debate. The point here is that the accounting system is not necessarily the best aggregator of information. The accounting system has a comparative advantage when it comes to providing information about the financial history of the firm. Market-related information might also reach the decision makers through other channels such as the pricing mechanism. In fact, the market mechanism is a powerful information aggregator, as suggested by Roll (1984). Hence, there is a horse race between the accounting system and the market mechanism in aggregating market information. When this aggregation is performed by the accounting system, the system will potentially be less efficient in providing historical financial information. As both types of information are in demand, there is a conflict. I return to this issue in Section IV.

II. THE INFORMATION PERSPECTIVE

Accounting provides information to decision makers. It produces numbers that have proven useful in pricing commodities and produces data that serve as inputs into the managerial performance evaluation exercise (Horngren et al. 2009). Thus, accounting provides information for the decision making and control purposes of the firm in general (Sunder 1997).

Bayes taught us how our beliefs are updated once new information arrives. The event of interest is called A, and the new piece of information is denoted y. The beliefs of the decision maker are encoded in the probability distribution, P. Bayes' formula shows that the updated probabilities reflect the relative likelihood of y when y has occurred:

$$P(A|y) = \frac{P(A \cap y)}{P(y)}. \tag{1}$$

More insight might be obtained when information is placed in a normal distribution framework. The value of the object in question is x, which is normally distributed with mean μ and variance σ^2. The new information signal is y, which is normal with mean x and variance σ_2^2. Consistent updating of the expectation of x given the information in y results in the following expression:

$x \sim N(\mu, \sigma^2)$ -value

$y \sim N(x, \sigma_2^2)$ -information

$$E[x|y] = \mu + \frac{\sigma^2}{\sigma^2 + \sigma_2^2}(y - \mu). \tag{2}$$

As anticipated, the updated expectation depends both on y and on the prior mean μ. The interesting observation is that the updated expectation of x also depends on the variances. The amount of updating depends upon the relative variance of the new information as compared to the prior variance. If the new source of information has very little variance—reliable information—then the updated value is equal to the signal. When the signal is of lower quality, the updating adjusts accordingly, as the magnitude of the updating is a decreasing function of the variance σ_2^2.

The implication of this for an accounting context can be shown in a simple setting, adapted from Christensen and Demski (2003a). Suppose that the firm has a lifetime of three periods and faces perfect capital markets with an interest rate of r. Initially, the owners invest $-CF_0 > 0$, generating uncertain cash flows over the following three periods of:

$$CF_1 = \bar{c}_1 + \varepsilon + \varepsilon_1$$

$$CF_2 = \bar{c}_2 + \theta\varepsilon + \varepsilon_2$$

$$CF_3 = \bar{c}_3 + \gamma\varepsilon + \varepsilon_3. \tag{3}$$

The expected value of all the error terms ε, ε_1, ε_2, and ε_3 is zero and the expected net present value of the investment is also zero. All error terms are independent, implying that the ε_i's are transitory earnings components, whereas ε has a permanent effect, as it affects all three periods. It is assumed that all cash is transferred to the owners as immediate dividends. Thus, the firm never holds cash.

The risk-neutral value of the firm is given at time 0 (value $= -CF_0$) and time 3 (value = 0). The interesting valuation questions are found at t = 1 and 2. Suppose only cash flows are observed by the market participants as time progresses. Given the efficiency of markets, the market valuation at time 1 will be:

$$E[PV_1|CF_1 = \bar{c}_1 + \varepsilon + \varepsilon_1] = E[CF_2|\varepsilon + \varepsilon_1](1+r)^{-1} + E[CF_3|\varepsilon + \varepsilon_1](1+r)^{-2} = (\bar{c}_2 + \theta E[\varepsilon|\varepsilon + \varepsilon_1])(1+r)^{-1} + (\bar{c}_3 + \gamma E[\varepsilon|\varepsilon + \varepsilon_1])(1+r)^{-2}. \tag{4}$$

The error terms provide the information for Bayesian updating. Furthermore, the autocorrelation of the error terms is what makes the valuation exercise interesting. Only the permanent component of the error term enters the value calculation.

The accounting system will also reflect the first period error term, but in a different way. Accounting is governed by accounting conventions and recognition rules. Under historical cost accounting, the accounting valuation will start from the book value (= acquisition value) and then use periodic depreciation to adjust the value. The depreciation charge is predetermined as δ_1. The first period accounting value is then $A_1 = A_0 - \delta_1$ and the period 1 accounting income is $I_1 = CF_1 - \delta_1$. The accounting income reflects the error terms. Since the accounting system reflects the same information as the market value, it is possible to calculate the market value using

accounting data. Contrary to the accounting income, the accounting valuation contains no information about the error term. The expected future residual income measures will counterbalance the error made in the accounting valuation, as in Feltham and Ohlson (1995). The calculation is as follows:

$$E[PV_1|\varepsilon+\varepsilon_1]=A_1+E[I_2-rA_1|\varepsilon+\varepsilon_1](1+r)^{-1}+E[I_3-rA_2|\varepsilon+\varepsilon_1](1+r)^{-2}. \quad (5)$$

The general point here is that the accounting valuation seldom coincides with the market valuation. The recognition rules of accounting dictate a more conservative information processing compared to the market valuation. When pricing the securities in a risk-free market, the market participants use their information to estimate the mean value and are not constrained by accounting conventions.

When the market valuation and the accounting valuation are based upon the same information set, it is possible to find the market valuation from the accounting valuation, making it important to extract the information about the permanent error term from the reported income. This analysis suggests that an important role of the accounting system is to provide information about the error terms, and yet how to fulfill that role is not predominant in the accounting literature.

When the market valuation is based on more information than the information contained in the accounting reports, it becomes more difficult to analyze the role of the accounting information. The errors are important, as suggested by Bayesian updating, but the question remains as to whether the accounting information should be confined to transaction information or whether market information should be part of the accounting system. I will return to this question in Section IV, after considering the effects of nonlinearities in accounting in Section III below.

III. NONLINEARITIES IN ACCOUNTING—COSTING

Costing is an old art of accountants and the ABC literature has given rise to a new interest (cf. Christensen and Demski 1995). One of the important insights to be gained from the activity-based accounting system is that the accounting system is improved when the accounting structure reflects the structure of production. In this way, ABC is better able to produce product costs that are in line with the underlying economic cost. The basic activity-based construction of product cost is a two-stage linear cost allocation. The errors of ABC have been the subject of recent research such as Cardinaels and Labro (2008). The decision problem regarding production plans, make-or-buy, or acquisitions might take on many disguises, but according to economics, marginal cost or incremental cost is the statistic that guides optimal decisions when confronted with the marginal revenue. Furthermore, cost functions are often nonlinear, as increasing or decreasing returns to scale are commonplace, such that nonlinearity is often cited as the reason for a particular size of an operation or even the existence of multi-product firms.

Accounting cost usually assumes a linear cost function. The accounting construction of unit cost is a prominent example of this. When the cost function is nonlinear, an endogenous error in the accounting system is the result.

Consider a facility that produces several products in quantities described by the vector q, as in Christensen and Demski (1997). Assume three indirect cost pools named after the production function: f, g, and h. Cost pool f uses the inputs x_1 and x_5 to produce the products in the quantity q. The products use the output from the cost pool in a Leontief technology described by A_1. This means that there is a linear use of the resources produced by f. In this way, A_1 illustrates the cost driver for pool one, as each unit of the first product uses the number represented by the first element of A_1 of the resource f. Cost pool g is similar to cost pool f. Cost pool h is an intermediate cost pool. The output of h is also Leontief, as it produces the inputs x_5 and x_6, which are used as inputs in f and g, respectively. The production function h uses the inputs x_3 and x_4. The production functions f, g, and h describe the transformation process of inputs to outputs for each of the cost

pools. In addition to these cost pools, production gives rise to labor cost, L, and material cost, M. These two cost categories are linear by default.

The cost function of the operation is derived from the following optimization:

$$c(q) = \text{Min}\left[Lq + Mq + \sum_{i=1}^{4} p_i x_i \right]$$

s.t.

$$A_1 q \leq f(x_1, x_5)$$

$$A_2 q \leq g(x_2, x_6)$$

$$x_5 + x_6 \leq h(x_3, x_4). \tag{6}$$

The marginal cost is the relevant cost statistic for incremental decisions. The Lagrange multipliers on the constraints are useful in characterizing the marginal cost:

$$MC_i(q) = L_i + M_i + \lambda_1(q)A_{1i} + \lambda_2(q)A_{2i}. \tag{7}$$

The Leontief coefficients fully capture the differences of the products' marginal costs, and there is a linear relationship between the marginal costs of the products. It is noteworthy that the intermediate cost pool is not present in this marginal cost calculation, though it is accounted for indirectly through Lagrange multipliers of cost pools f and g. The marginal cost is not necessarily linear in the products q. That depends upon the production functions f, g, and h, as the Lagrange multipliers are not necessarily constant.

The costs are accumulated in the three obvious cost pools. The costs are exactly as suggested by the cost functions and the production functions as in Equation (6). There are no error terms and there are no inefficiencies. Now apply cost calculations of the ABC type to find the product costs. First, the intermediate cost pool is allocated to the indirect cost pools f and g using the available cost drivers describing the activity of the intermediate cost pool. Second, the total cost in the indirect cost pools is allocated to the cost objects using the Leontief coefficients of the production functions. These will form the natural and available cost drivers. The calculations are summarized in Equation (8):

$$C_h = p_3 x_3 + p_4 x_4$$

$$C_f = \frac{p_1 x_1 + \dfrac{x_5}{x_5 + x_6} C_h}{A_1 q}$$

$$C_g = \frac{p_2 x_2 + \dfrac{x_6}{x_5 + x_6} C_h}{A_2 q}$$

$$C_i = L_i + M_i + C_f A_{1i} + C_g A_{2i}. \tag{8}$$

It is noteworthy that the structure of the ABC calculation is identical to the structure of the marginal cost calculation. The direct cost components are identical and the Leontief coefficients A_{1i} and A_{2i} are also found in the ABC calculation. Thus, the differences among the products with

respect to cost structure are precisely reflected in the ABC system, due to the Leontief structure of the production function. Furthermore, any discrepancy between ABC cost and marginal cost must be found in the Lagrange multipliers, which represent the marginal costs of the indirect cost pools. As noted, the multipliers indirectly account for the cost of the intermediate cost pool as in the ABC system. When the production functions f, g, and h are all linear, the ABC approach results in the marginal cost for all of the products simultaneously.

When one or more of the production functions show increasing returns to scale, the ABC system will estimate the marginal cost too high, whereas the opposite is the case for decreasing returns to scale.

The activity-based accounting system was introduced as an improvement of the traditional accounting system, particularly for firms that produce multiple products. In comparison, the traditional accounting system is a single cost pool system using direct labor as the cost driver. The product cost calculation runs according to the following pattern:

$$C_L = \frac{p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4}{Lq}$$

$$C_i = L_i(1 + C_L) + M_i. \tag{9}$$

This costing system is simpler than the ABC system, as it has only one independent parameter to describe the differences of the product cost of the products in q. The input of labor, L, is the only source of assigning overhead costs to the products.

However, as the ABC system is also an approximation, it is worthwhile to run the horse race between these two systems, as in Christensen and Demski (1997). The actual marginal cost of the products is used to evaluate the performance of the two accounting systems. The performance of an accounting system is measured by the mean squared error of the unit cost reported by the accounting system. Christensen and Demski (1997) found that the traditional accounting system based solely on labor input for some combinations of the production functions outperformed the more refined ABC system. The picture is more pronounced when the individual products are evaluated. For some products, one accounting system performs best, whereas for other products, the other accounting system performs best. There is no uniform winner.

This finding is not restricted to nonlinear production functions. Datar and Gupta (1994) found similar results when they studied the combined effect of an erroneous choice of cost driver and different levels of aggregation. They found that the two factors interact, such that repairing the system on one dimension does not necessarily lead to an overall improvement. They also used the Euclidean distance measure to determine the performance of the accounting systems. Labro and Vanhoucke (2007) took the analysis one step further by performing a large simulation analysis. The underlying "true" economics of the firm is linear and with a multi-layer system like an ABC system. Their setting included a large number of cost pools that are allocated to activity cost pools and further to cost objects. All the allocations are linear, using pre-specified cost drivers. In their simulation, Labro and Vanhoucke (2007) varied the number of cost pools both at the resource cost pool level and the activity cost pool level. They also varied the cost drivers used in the cost allocations. Results supported the findings of Datar and Gupta (1994), revealing interaction effects between errors. An incremental refinement of the accounting system does not always lead to a superior system.

The same pattern of results arises when the accounting system is used to induce efficient behavior. The economic relationships in the division are as described in Equation (6). It is assumed that part of the production from this division is sold for profit in the market, and part of the

production is delivered to other divisions in return for cost reimbursement. The cost reimbursement is based on the realized cost and observed drivers, as measured by the accounting system.

The evaluation of the division is based upon divisional profit. Inefficient decisions result in excessive costs, but perhaps divisional profit will increase. The question is whether the division is able to pass more than the excessive cost on to the other division through clever but inefficient decision making. The control system prevents this from happening to some extent. Christensen and Demski (2003b) found that, when the technology exhibits constant returns to scale, the ABC-type accounting system will always induce efficiency. However, if the indirect cost pools show increasing returns to scale or if the intermediate cost pool has non-constant returns to scale, then it is possible for the division to profit from using resources inefficiently. The inefficiency is passed on to the other division via the cost allocation. The scheme relies heavily on substitution of factors of production. The scheme allows the division to employ the more productive resources for the products sold for profit, whereas the less productive inputs are used for internal transfers. There is nothing in the system itself which limits the amount of inefficiency. Such limits must be found in the control system.

When the possibility of such inefficiencies exists, a simpler traditional accounting system might be better at promoting efficiency. Christensen and Demski (2003b) analyzed the horse race between a labor-cost-based allocation system and an activity-based accounting system. The production technology suggests an activity-based system. The inefficiencies of an ABC system are confined to the indirect cost categories.

If labor is used to allocate cost and if it is possible to use too much labor for the internal division, then the internal division absorbs a larger share of the indirect cost. The inefficiency is confined to the direct labor and also in this case the control system limits the inefficiency.

The net result is that the winner of the contest is determined by the control system. Are the controls better at limiting the inefficiencies in the labor domain than in the indirect cost domain? Accounting provides a mapping and illustration of the economics of the firm. The accounting model has many parameters, and selecting the model just to mirror the structure of the firm is too simplistic. The underlying economics play a crucial role. The other dimension is the control system, which might result in one seemingly inferior accounting model winning the horse race against a more refined accounting model. Unlike accounting, the world is hardly ever linear, and the analysis of accounting has to take that reality into consideration.

IV. ALTERNATIVE INFORMATION SOURCES

Traditionally, accounting has been based upon historical cost and revenue measures. Only historical information or realized cash flows could penetrate the recognition rules of accounting. Accruals also have a historical orientation, but contain certain forward-looking elements. The valuation of material assets starts out with historical cost and is then depreciated according to a predetermined scheme. The scheme is determined at the time of acquisition. Only hard evidence will make it possible to deviate from the scheme. Recently, there has been a trend toward substituting the historical valuation of assets with a fair market valuation, in which the current market value of the asset is used in the accounting valuation of the asset.

Fair-value accounting changes recognition rules to include market information in the accounts to a larger extent. The implication is that the aggregation within the accounting system reduces the information content of the accounting statement with respect to historical information. Consequently, fair-value accounting poses a new horse race concerning introduction of alternative information sources into the accounting system. One system would be the historical accounting system, which is transaction based, and opposed to that is a fair-value accounting system, which allows market information to be aggregated into the accounting system. Market information is available in the market by definition and market participants also have access to this type of

information. Therefore, the horse race could also be coined into a question of who is best at performing the aggregation of market-wide information with historical cost information. The contestants for this competition are the market participants or the accountants. The analysis of this problem follows Christensen and Frimor (2007).

Inherently, the value of a firm is uncertain. To keep the presentation simple, suppose that the uncertainty consists of two components: one that summarizes the internal factors and the other that summarizes the external factors of the firm.

$$V = \varepsilon_A + \varepsilon_E. \tag{10}$$

It is assumed that the firm observes the first error term and only observes the second error term with error. Thus, the firm and its auditor observe:

$$y_A = \varepsilon_A, \quad y_F = \varepsilon_E + \varepsilon_{AU}. \tag{11}$$

The firm and its accountant are confined to a single aggregated report stating the expected value of the firm conditional on the information that is allowed by the recognition rules. There are two possibilities for the recognition rules that lead to two reporting options. One is a historical cost model in which the recognition rule only allows the internal factors of the firm into the accounting valuation, V_H. The alternative set of recognition rules allows the accountant to aggregate the observed external factors into the reported value of the firm. This set of recognition rules brings market observations into the accounting valuation, and in this sense it has the flavor of fair-value accounting, V_F. In both cases, the accountant is disclosing an unbiased report subject to his/her information constraint. The report published by the firm under the two regimes can be calculated assuming that all information variables are normally distributed with zero means:

$$V_H = E(V|y_A) = y_A$$

$$V_F = E(V|y_A, y_E) = y_A + \frac{\sigma_E^2}{\sigma_E^2 + \sigma_{AU}^2} y_E. \tag{12}$$

Investors observe the accounting report prior to trading. Each investor also makes his/her own private observation of the external factor, and the error term for each investor is personal but follows the same distribution. To simplify, it is assumed that all the error terms of the investors are uncorrelated. In addition, the investors observe the price, P, at which the stock of the firm is traded. Thus, the investors observe:

$$y_{E,i} = \varepsilon_E + \varepsilon_i, P, V_i. \tag{13}$$

The security is traded in the open market, and the price is formed when the investors meet and decide on their individual trading. There is uncertainty in the market, as all individual endowments are stochastic and assumed normal with variance σ_z^2. This structure leaves the model tractable and prevents prices from being too informative. Supposedly, the investors participate in an anonymous trading mechanism transforming the individual bids and forming a price. The investors only disclose their private information through the pricing mechanism, and there is no direct information exchange among the investors. The investors are rational, and when entering into the pricing game, they anticipate the reaction of their co-traders. The market clearing price reflects this conjecture, such that the resulting price is a rational expectations equilibrium. The model used for this analysis stems from Hellwig (1980).

The equilibrium price for this mechanism reflects the market uncertainty, σ_z^2, the private information of the investors, and the information content in the accounting report.

The important question is how to evaluate the two accounting options. One method is to analyze how the accounting valuation is aligned with the market valuation. The relevant measure for this is the squared difference between the accounting valuation and the market value. This difference is often used in empirical research, as both measures are readily available (Schipper and Vincent 2003). This measure would clearly favor the market-based accounting valuation as the accounting valuation includes the second error term.

An evaluation of the accounting system is closely linked to the ability of the accounting system to provide useful information for decision making. The final outcome of the decisions is what matters to the investor.

The market value might be a good proxy for the final outcome. At the same time, the market value works as an information aggregator, compounding all information available to investors, including the private information and the accounting information. The aggregator role is not highlighted when only the difference between the accounting value and the market value is evaluated. Using the market value as a measure provides only a partial view of the relative performance of the two accounting systems. Consequently, market value is not necessarily a good measure of the merits of the accounting system. Furthermore, another measure of the merits of the two accounting options would be the squared difference between the price of the asset and the final payout from the investment. This is a measure of the aggregate informativeness of the accounting system. It is difficult to measure the final payout empirically, as the final dividend only materializes with a considerable time lag for most investments. In the present setting it is possible to evaluate this measure analytically. The measure reflects the ability of the price to reflect the information available to the market participants. It is of no relevance to the measure how the information is transferred to the price. As a result, it provides a measure of who is best at aggregating the information: the accountant or the market.

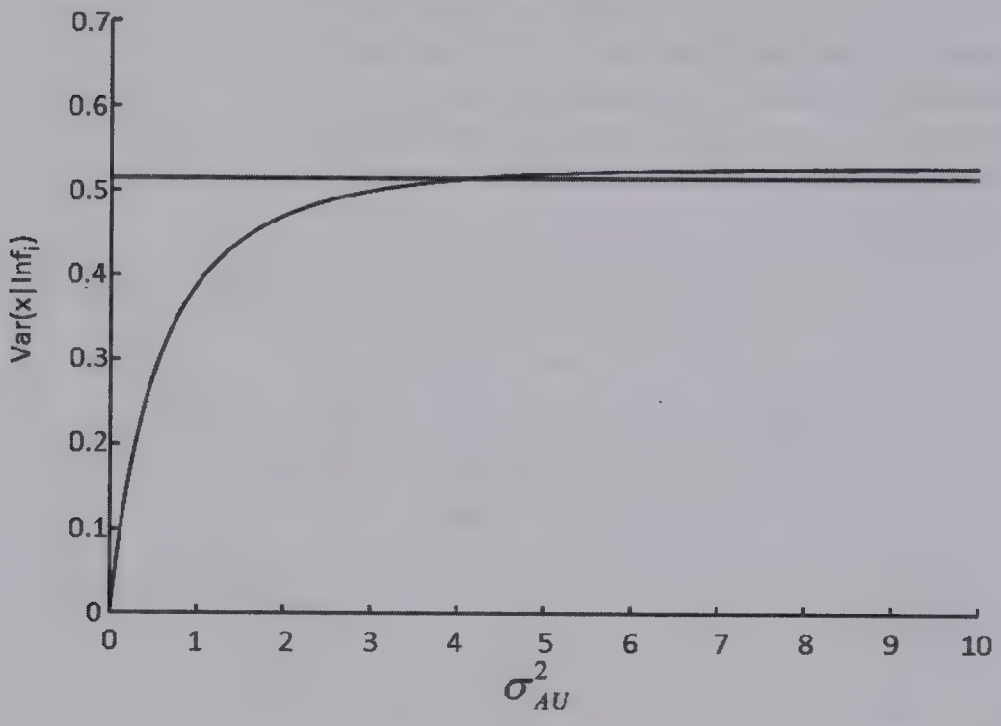
It is not easy to get a closed form solution to show exactly which accounting system is optimal, but it is possible to calculate the value of the valuation error as a variance, Var(.). The valuation error is calculated for the two accounting systems. The historical accounting system is independent of the accountant's ability to measure the market component. The accounting valuation of the fair-value accounting system depends on the quality of the accountant. Consequently, the variance is calculated as a function of the ability of the accountant to estimate the "market" information for the fair value accounting system—i.e., as a function of the variance σ^2_{AU}. For example, if the auditor is very good at estimating the market factor ε_E, then the variance σ^2_{AU} is very small.

In Figure 1, the performances of the two accounting systems are calculated when market uncertainty is large—that is, when the variance of the initial wealth of the investors is large. In that case, it is expected that the market price will carry little information from the individual investors and, consequently, even less-able accountants will be able to provide useful market information to the price mechanism. The graph for the fair value accounting system shows a smaller variance, even for relatively large σ^2_{AU}.

The calculations are repeated for smaller market uncertainty in Figure 2. In that case, it is expected that the market price will reflect the investors' private information to a large extent, and as a result there is a greater demand for accounting information to supplement the information that is already available in the market. Therefore, it is expected that the historical accounting system will outperform the fair-value accounting system. A comparison of the graphs in Figure 2 reveals that this is indeed the case. The two graphs intersect at $\sigma^2_{AU} = 0.02$, implying that the accountant should outperform the average investor 50 times when it comes to estimating the market influence on the performance of the firm.

These findings support the view that the optimal accounting system depends on the finer details. In some cases, it is optimal to aggregate other information into the accounting system, and

FIGURE 1
High Market Uncertainty



in other cases, it is optimal to let the market do the aggregation. The other message is that the accounting system is not the only system available and it is a serious error to pretend that the accounting system should be the only information system.

V. CONCLUDING REMARKS

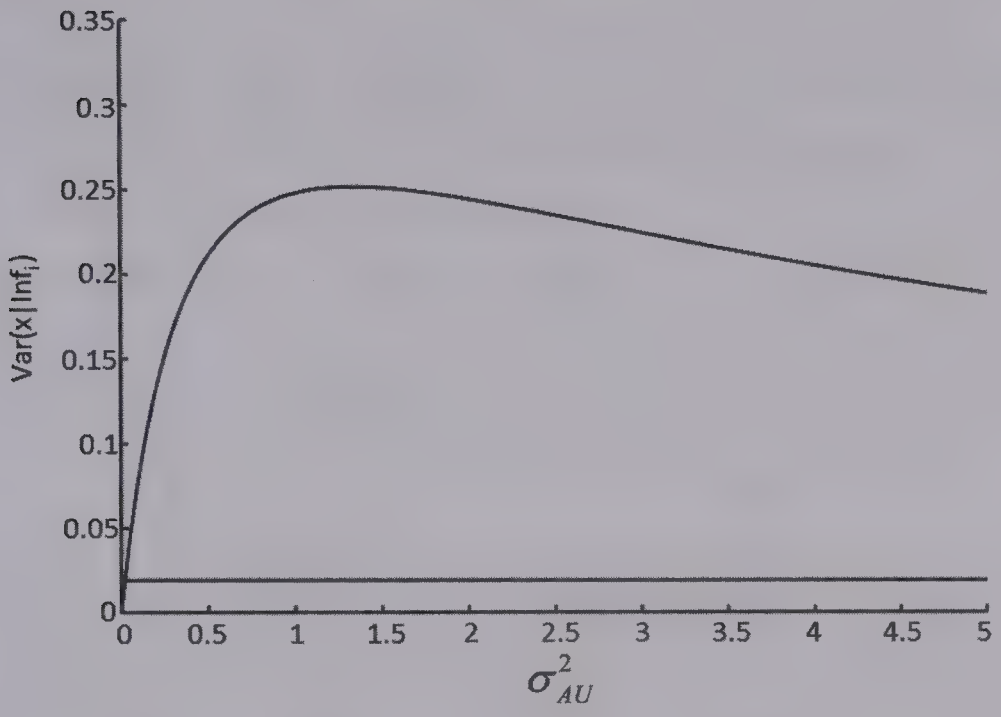
The common theme of this discussion has been the errors of accounting systems. Errors are important and often neglected when the design of accounting systems is evaluated. Accounting carries information, as is often recognized, but the full impact of that statement is not endogenized in the analysis (Demski 2004).

First, the information perspective calls for using the accounting information to update our expectation of the future events of the firm, cost, cash flows, etc. The important element of this is what we learn from the accounting statement and what we knew beforehand. The quality of both pieces of information is important for consistent updating of beliefs in a manner consistent with Bayes' theorem.

The accounting mapping is a linear construction, using a linear aggregation of transactions in the accumulation accounts, in the allocation of cost and benefits to cost objects, and in the presentation in the balance sheet. Yet, the world is hardly linear. Economies of scale and scope are found everywhere and are often the reason for the existence and size of firms. Thus, the assumed linearity in accounting results in structural errors. There are many reasons for sticking to the linearity assumption, as estimating the "correct" functional form of some accounts is an impossible task. The result is that the errors persist, and the system designer can counter this through the allocation mechanisms employed. Understanding the error structure is important.

Finally, the accounting system is not the only information system in the world. Other and perhaps more timely information sources exist. The accounting system must acknowledge this fact

FIGURE 2
Low Market Uncertainty



and act accordingly. The comparative advantage of the accounting system must be considered, with some tasks better left for other information sources. The debate on fair-value accounting provides an example of this reasoning. The accounting system has a non-disputed advantage when it comes to reporting of firm specific information. Aggregating market information into the accounting system might or might not be a good idea. The basic question here is who is best at performing the aggregation of the market information. The market mechanism is a powerful information aggregator, and it might be better suited for aggregating firm-specific information with market information.

REFERENCES

Beaver, W. 1998. *Financial Reporting: An Accounting Revolution*. Upper Saddle River, NJ: Prentice Hall.

Cardinaels, E., and E. Labro. 2008. On the determinants of measurement error in time-driven costing. *The Accounting Review* 83 (4): 735–756.

Christensen, J., and J. Demski. 1995. The classical foundations of modern costing. *Management Accounting Research* 6 (1): 13–32.

———, and ———. 1997. Product costing in the presence of endogenous subcost functions. *Review of Accounting Studies* 2 (1): 65–87.

———, and ———. 2003a. *Accounting Theory. An Information Content Perspective*. New York, NY: McGraw-Hill/Irwin.

———, and ———. 2003b. Factor choice distortion under cost based reimbursement. *Journal of Management Accounting Research* 15: 145–160.

———, and H. Frimor. 2007. Fair value, accounting aggregation and multiple sources of information. In *Essays on Accounting Theory in Honour of Joel S. Demski*, edited by Antle, R., P. J. Liang, and F. Gjesdal, 35–51. New York, NY: Kluwer.

Datar, S., and M. Gupta. 1994. Aggregation, specification and measurement error in product costing. *The Accounting Review* 69 (4): 567–591.

Demski, J. S. 2004. Endogenous expectations. *The Accounting Review* 79 (2): 519–539.

Feltham, G., and J. Ohlson. 1995. Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research* 11 (2): 689–731.

Gjesdal, F. 1981. Accounting for stewardship. *Journal of Accounting Research* 19 (1): 208–231.

Hellwig, M. 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22 (3): 477–498.

Horngren, C., S. Datar, G. Foster, M. Rajan, and C. Ittner. 2009. *Cost Accounting: A Managerial Emphasis*. New York, NY: Prentice Hall.

Labro, E., and M. Vanhoucke. 2007. A simulation analysis of interactions among errors in costing systems. *The Accounting Review* 82 (4): 939–962.

Roll, R. 1984. Orange juice and weather. *The American Economic Review* 74 (5): 861–880.

Schipper, K., and L. Vincent. 2003. Earnings quality. *Accounting Horizons* 17 (Supplement): 97–110.

Sunder, S. 1997. *Theory of Accounting and Control*. Cincinnati, OH: South-Western.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 1839–1860

American Accounting Association
DOI: 10.2308/accr.2010.85.6.1839

Informativeness, Incentive Compensation, and the Choice of Inventory Buffer

Stanley Baiman
Serguei Netessine
University of Pennsylvania

Richard Saouma
University of California, Los Angeles

ABSTRACT: Previous research in management accounting and economics has noted the potential for complementarities between the firm's performance measurement system and its other organizational design choices. We add to this literature by studying how the informativeness and incentive properties of a performance metric can be influenced by one particular organizational design choice—the size of the firm's inventory buffers. We model a manufacturing setting in which an agent manages a workstation that processes intermediate units. As intermediate units arrive, they are stored in an inventory buffer until the agent can process them. The buffer can hold a maximum number of intermediate units—its *buffer size*. The agent is compensated on the basis of his workstation's throughput. We characterize the conditions under which reducing the inventory buffer enhances/degrades the informativeness of the performance metric and, hence, mitigates/exacerbates the agent's incentive problem.

Keywords: *agency theory; inventory buffers; incentive complementarities; organizational design; slack.*

I. INTRODUCTION

Most of the agency theory work in managerial accounting has studied the relation between performance metric properties (e.g., informativeness, precision, congruity) and optimal incentive compensation, holding fixed the firm's other organizational design decisions (e.g., hierarchical structure, job assignment, production technology). However, by holding fixed these other organizational design choices, one cannot study whether and how they affect the

This paper has benefited from the comments of workshop participants at the University of Wisconsin–Madison, The University of Iowa, University of Houston 2007 Accounting Symposium, 5th Accounting Research Workshop (Fribourg, Switzerland), The University of Chicago, University of Pennsylvania, Tel Aviv University, Hebrew University (Jerusalem), Norwegian School of Economics and Business Administration, and Catholic University of Portugal, especially those of R. Balakrishnan, P. Berger, R. Cazier, D. DeJong, J. Gerakos, F. Gjesdal, J. Glover, T. Hemmer, E. M. Matsumura, M. Penno, S. Radhakrishnan, K. Sivaramakrishnan, and J. Stecher. We are particularly indebted to Jack Hughes, Cathy Schrand, the editor, and two anonymous reviewers for many helpful suggestions, and to Romanos Malikiosis for his outstanding research assistance.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: May 2008
Accepted: May 2010
Published Online: November 2010

performance metric properties. As Hemmer (1998, 321–322) notes, "the value of a performance measure is determined not simply by its congruity and precision but by its influence on the optimal organizational design ... Much of the recent theoretical accounting literature ... has largely ignored complementarities between performance measures and organizational design."

In this study, we expand on Hemmer's (1998) observation by studying how the informativeness and incentive properties of a performance metric are influenced by one of the firm's organizational design choices—the size of its inventory buffers. The introduction of just-in-time (JIT) and, more generally, lean manufacturing has led to an increased emphasis on controlling and reducing inventory levels. Managerial accounting has responded to this reduction in inventory with new costing techniques such as backflush costing (Horngren et al. 2008). However, we show that the choice of inventory buffers has a more subtle effect on the design of the managerial accounting system, in that it affects the informativeness of performance metrics produced by the managerial accounting system.

We model a manufacturing setting in which an agent expends effort to set up or program a workstation that processes intermediate units. As intermediate units arrive, they are stored in an inventory buffer until the agent can process them. The buffer can hold a maximum number of intermediate units, b—its *buffer size*. The buffer size represents the workstation's maximum order backlog. If an intermediate unit arrives while the buffer is full, then the unit is diverted and the agent loses the opportunity to process it. If the workstation is ready to process another unit but the buffer is empty, then it remains idle until the next intermediate unit enters the buffer. While we frame the analysis in terms of a manufacturing setting, the analysis is also applicable to service operations such as call centers that can have a maximum number of calls awaiting service.

We assume that the agent is compensated on the basis of his realized throughput; i.e., the *output per unit of time* of his workstation.[1] We show that the principal's choice of buffer size indirectly affects the effectiveness with which throughput reflects the agent's chosen effort and consequently affects the optimal compensation weight placed on throughput. The buffer size affects both the probability that the workstation's buffer will be full when an intermediate unit arrives, causing it to lose the opportunity to complete that unit (referred to as *blocking*), and the probability that the buffer will be empty when the workstation is free, causing it to forgo further production until another intermediate unit arrives (referred to as *starving*). Both probabilities affect the marginal productivity of the agent's setup effort with respect to throughput (his performance metric) and, hence, the extent to which that effort is reflected in throughput.

Next we discuss the relevant literature on organizational design and incentives. Section II introduces the model and Section III presents the results. Section IV discusses extensions and concludes.

A number of articles in both the accounting and economics literatures have considered the effect of organizational design choice on the design of optimal incentives. Among the issues considered are: job design (Riordan and Sappington 1987; Holmstrom and Milgrom 1991; Balakrishnan et al. 1998); the choice of production systems and technologies (Hemmer 1995, 1998; Milgrom and Roberts 1995); the hierarchical structuring of work (Melumad et al. 1995); routing schemes for product rework (Lu et al. 2009); and production bottlenecks (Datar and Rajan 1995; Gietzmann and Hemmer 2002).[2]

---

[1] Realized throughput is measured as realized total output during the shift that the workstation has been operating divided by the length of the shift.

[2] Also somewhat related is Cremer (1995), which examines the incentive effects on both the buyer and supplier when the buyer goes to a zero-inventory system for the supplier's product. However, the two motivations are different in that Cremer (1995) is based on a double moral hazard model.

Similar to our model, Hemmer (1995, 1998) and Gietzmann and Hemmer (2002) examine how different workflow arrangements between agents affect the information available for contracting, and the incentives facing agents. Our work is distinct from these models that do not consider buffer size as one of the principal's choice variables. Nagar et al. (2009) examine the role of inventory buffers in agency problems, although unlike our model, in their setting buffers are filled by agents in order to signal private information and buffer size is not a choice variable. Alles et al. (1995) do focus on the effect that the choice of buffer size can have on the informativeness of performance metrics. The major difference between our work and theirs is that they do not formally model how buffers affect performance metrics, but instead assume a monotonic relation. In contrast, we formally model the inventory process and derive a non-monotonic relation between buffer size and the informativeness of the agent's performance metric.

II. MODEL AND INITIAL ANALYSIS

The model consists of a single-period and a firm comprised of a risk-neutral principal and agent. Contracting takes place at the start of the period when the agent is hired to set up a workstation to process incoming intermediate units (e.g., test completed computers or weld automotive chassis). The intermediate units arrive stochastically into the workstation's incoming inventory buffer at a commonly known, mean arrival rate of λ. The buffer has capacity $b \in \mathbb{N}$, allowing a maximum of $b - 1$ units to be held in front of the workstation while the workstation processes one unit. If an intermediate unit arrives and the buffer is not full, then the unit is added to the inventory. However, if the buffer is full, then blocking occurs and intermediate units cease arriving until there is space in the buffer, whereupon the intermediate units begin arriving at the same stochastic rate as before.[3]

Once parties agree to the contract, but before the workstation begins processing, the agent chooses his personally costly setup effort r, which affects the average rate at which his workstation can process the intermediate units. That effort is unobserved by the principal and is subject to moral hazard. The arriving intermediate units are sufficiently heterogeneous that the agent's effort does not perfectly control the rate at which the workstation processes the incoming intermediate units. Rather, the intermediate units are processed at a stochastic rate with mean $r(1 + h)$, where $h \in (-1,\infty)$ represents the productivity-enhancing resources the principal has allocated to the agent. Examples of the latter include: machine and tool upgrades, improved working conditions, training, etc. As the workstation finishes processing an intermediate unit, the agent reaches into the buffer for the next unit on which to work. If the buffer is empty, then starving occurs and the workstation is idle until the next intermediate unit arrives. Figure 1 provides a graphical description of the process.

The principal chooses both the agent's compensation scheme and the buffer size, b. We restrict our attention to linear compensation schedules for the agent consisting of a fixed payment, w_0, plus an incentive wage, w, based on realized throughput—the number of units processed per unit of time.[4,5]

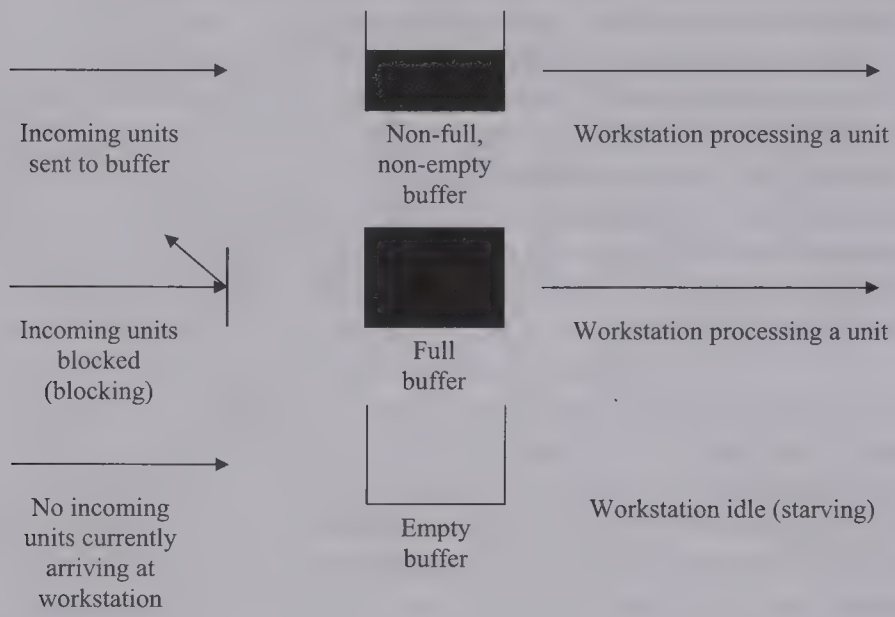
---

[3] Alternatively, one can interpret the model as one in which the intermediate units continue arriving while the inventory is blocked but are diverted elsewhere. What is important is that in both cases, blocking imposes an opportunity cost on the agent because his workstation loses the chance to process the intermediate units that might have been added to the buffer had it not been full.

[4] Given that the basic properties of queuing systems are stated in terms of rates (i.e., arrivals per unit of time, number of units processed per unit of time) we assume that the principal compensates the agent on throughput per unit of time rather than total throughput. Linear compensation schemes are frequently used in practice (Berg and Fast 1975; Hall et al. 2000) and are extensively used in the analytical literature (Feltham and Xie 1994; Alles et al. 1995; Hemmer 1995, 1998; Gietzmann and Hemmer 2002).

[5] In Section IV, we discuss the possibility of contracting on inventory levels.

FIGURE 1
The Effect of Inventory on the Production Process



In the top panel, the buffer is not full, incoming units are added to the buffer, and the workstation starts processing a unit from the buffer once it completes the current unit. In the middle panel, the buffer is full, incoming units are blocked from entering the buffer (blocking) until the workstation finishes processing its current unit and the agent withdraws the next intermediate unit from the buffer. In the bottom-most panel, the buffer is empty, implying that the workstation is idle (starved) until the next unit arrives.

Last, we assume that the marginal cost associated with each unit of buffer capacity is $c > 0$, the agent's marginal cost of effort is $a > 0$, and we normalize the marginal cost of the additional resources provided to the agent by the principal, h, at one dollar per unit of resource.

All notation and assumptions are as follows:

λ = The mean arrival rate of the intermediate units. The arrival rate is Poisson distributed, implying that the time between arrivals is distributed according to an exponential distribution with mean of λ^{-1}.[6]

r = The level of effort the agent exerts at private cost, ar.

h = The amount of productivity-enhancing resources allocated to the agent. The principal's out-of-pocket cost is h.

$r(1 + h)$ = The mean rate at which the agent's workstation processes the intermediate units. The actual processing time is distributed according to an exponential distribution.

b = The maximum number of units the station can simultaneously store (buffer size). The principal incurs the associated capacity cost, cb, each period.

[6] For an exponential distribution, the same parameter that represents the mean also represents the variance. For simplicity we refer to the choice of this parameter as the choice of the mean arrival rate.

K = The steady-state mean throughput per period, with $\tilde{K}$ denoting the realized throughput.

w, w_0 = The incentive payment made to the agent per unit completed per period, and the fixed salary payment, respectively.

Note that in our model there are two sources of noise that make it difficult for the principal to infer the agent's effort from throughput. As is common in agency problems, the relation between the agent's effort and the time required for the workstation to process an intermediate unit is stochastic. In addition, the rate at which units become available to be processed is also stochastic. Thus, even if the relation between the agent's effort and the time required to process an intermediate unit were deterministic, instantaneous throughput would still be stochastic in the agent's effort because of the stochastic arrival rate of the intermediate units. Placing an inventory buffer in front of the agent's workstation is one way of dampening out the effect of the variation in arrivals; however, as we will show, such dampening can also reduce the informativeness of throughput with respect to the agent's effort.

Our assumptions regarding arrival and processing rates follow the standard *M/M/1/b* queuing model with finite buffers used in the operations management literature.[7] Note that while we are modeling a single-period in that the agent chooses his setup effort only once at the start of the period, we analyze the stochastic evolution of the workstation's throughput over the entire period. The *transient* behavior of throughput in queuing systems with finite buffers cannot be described analytically, although the *steady-state* behavior can be described in closed form.[8] As a result, the operations management literature typically uses simulations to analyze the transient behavior of queuing models with finite buffers and closed-form analysis to examine steady-state behavior. In contrast, agency models have emphasized rewarding transient behavior, but simplify the production process (e.g., using the LEN model) to achieve closed-form solutions.[9] In order to incorporate stochastic intermediate unit arrival rates, stochastic processing rates, and finite inventory buffers in our analysis, we focus on the problem's *steady-state*. That is, we assume that the principal is interested in maximizing her *steady-state* expected profit and that the agent selects his effort to maximize his *steady-state* expected utility.[10] This is a reasonable approximation if, as we assume, the production process reaches steady-state fairly quickly. Our assumption that the agent's effort is expended in setting up the workstation before production begins, rather than managing the production process as it evolves, is also consistent with our emphasis on the steady-state behavior of the workstation. An example of our modeled setting is a robotic line that welds automotive chassis. The agent's main task is to program the spot-welding robots. This task involves identifying the optimal position of the chassis, identifying the optimal position of the robotic arms, and making sure that robot paths do not cross each other or the chassis itself (so that nothing is damaged in the process) while ensuring that welding is done in the least possible number of steps. All of this planning and programming setup is done once by the agent at the beginning of the shift and thereafter, as long as the workstation is in control, the agent does not need to intervene. Clearly, the actions of the agent in programming the robots affect the time it takes to weld the chassis (on average), but the agent does not reprogram the robot once production has begun.

---

[7] The *M/M/1/b* model is the most widely accepted modeling approach for production systems with finite buffers (Berkley 1992; Groenevelt 1993; Balsamo et al. 2001).

[8] The process attains steady-state when the probability distributions over inventory levels and throughput are no longer functions of the starting conditions, and no longer vary over time. That is, while the realized inventory level and throughput at any moment are still uncertain, the probability distributions that describe them do not vary over time.

[9] Queuing models similar to ours have been used to analyze the role of cost allocation in resolving congestion problems (Radhakrishnan and Balachandran 1995, 2004; Balachandran and Radhakrishnan 1996).

[10] Lu et al. (2009) restrict their analyses to steady-state for the same reason.

The workstation's realized output per unit of time or throughput, $\tilde{K}$, is stochastic, though the *mean* steady-state throughput, K, can be represented as (Hopp and Spearman 2001):

$$K(b,r) = \lambda(1 - p_b) = \lambda \frac{1 - (r(1+h)/\lambda)^{-b}}{1 - (r(1+h)/\lambda)^{-(b+1)}}, \quad (1)$$

where p_b is the steady-state probability of the buffer being full; i.e., the steady-state probability of blocking. The workstation's steady-state mean throughput is thus equal to the unblocked mean arrival rate of intermediate units, λ, times the steady-state probability that the buffer can accept an additional unit when it arrives, $1 - p_b$. In equilibrium, the workstation's mean steady-state throughput is increasing in: r, the agent's effort; h, the level of resources allocated to the agent; b, the size of the input buffer; and λ, the mean arrival rate of the intermediate units. Equation (1) illustrates that the agent's effort affects throughput via its effect on the probability of blocking, or equivalently, the probability of starving.[11]

Note that the mean arrival rate of incoming intermediate units and therefore the maximum mean throughput is λ. However, if the principal induced the agent to choose $r = \lambda/(1 + h)$ so that the workstation was set to process the intermediate units at the same average rate at which they arrived, then the instantaneous probability of starving the workstation would be $p_s = \frac{1}{1+b} > 0$, which in turn would limit the workstation's mean steady-state throughput to $K = \frac{\lambda b}{1+b} < \lambda$. That is, inducing the agent to set the workstation to process the intermediate units at the same average rate at which they arrive would significantly limit steady-state throughput and induce a non-trivial probability of starving. Thus, the principal has a natural incentive to induce the agent to set the workstation to process units at an average rate that is strictly greater than the average rate at which the intermediate units arrive, as commonly assumed in the operations management literature (Hopp and Spearman 2001).

Assume that the principal earns R dollars per unit of throughput. The principal chooses the set of decision variables, (w,w_0,b), and a target level of agent setup effort, r^*, which maximize her steady-state objective function:

$$Max_{(w,w_0,b,r^*)} RK(b,r^*) - wK(b,r^*) - w_0 - cb - h.$$

Given (w,w_0,b), the agent chooses his setup effort, r, to maximize his steady-state objective function $max_r\, wK(b,r) + w_0 - ar$. Varying the size of the buffer, b, has three effects on the principal's problem: (1) it incurs a carrying cost of c per unit of buffer, (2) it affects the principal's expected revenue through its direct effect on steady-state throughput, $K(b,r)$, and (3) it indirectly affects the agent's effort incentive through its effect on the expected steady-state throughput, $K(b,r)$. Effort and buffers are substitutes in throughput and, therefore, the principal's choice of b affects the wage, w, required to induce the agent to work. Thus, the latter two effects of the principal's choice of b directly influence the effort that the principal wants to induce the agent to take and the cost of inducing that effort.

As noted earlier, we are interested in studying the *incentive* implications of the principal's choice of the buffer, b. To isolate that effect, we will analyze the problem in which the principal wants to minimize her expected steady-state cost of inducing the agent to take a *particular* effort, r^*. This allows us to hold constant the principal's desire to choose b so as to directly affect revenue and the carrying cost associated with buffers, and instead concentrate on her desire to

[11] The two probabilities are directly related in that the probability of starving is $p_s = 1 - \frac{\lambda}{r(1+h)}(1 - p_b)$, allowing (1) to be rewritten as $K(b,r) = r(1 + h)(1 - p_s)$. Note that a change in the buffer affects the probability of starving in the same direction (i.e., increase/decrease) as its effect on the probability of blocking.

choose b to affect the cost of inducing effort from the agent. Thus, we take a Grossman and Hart (1983) approach to the principal's expected profit maximization problem. This is without loss of generality because any incentive effects of inventory buffer size found in this analysis will also appear in the principal's full profit-maximization problem.[12] Consequently, our subsequent analysis is analogous to the first step in a Grossman and Hart (1983) analysis—minimizing the principal's expected steady-state cost of inducing effort r^* from the agent.

The problem can be written as:

$$Min_{\{w,w_0,b,r^*\}} wK(b,h,r^*) + w_0 + h \qquad \text{(Program 1)}$$

s.t.

$$w\tilde{K} + w_0 - ar^* \geq 0 \ \ \forall \tilde{K} \qquad \text{(LL)}$$

$$r^* \in \arg\max_r (wK(b,r) + w_0 - ar). \qquad \text{(IC)}$$

We assume that the labor market is such that the principal must leave the agent with a minimum level of utility regardless of realized throughput (constraint (LL)).[13] The incentive compatibility constraint, (IC), ensures that the agent will select the principal's desired level of effort r^*. In the First-Best case, in which the agent's effort is not subject to moral hazard, (IC) can be ignored and, given that the minimum possible realized throughput is $\tilde{K} = 0$, the limited liability constraint, (LL), implies that $w_0 = ar^*$ and $w = 0$. Hence, the size of the buffer b plays no role in the agent's compensation in the First-Best solution. In the Second-Best setting, (LL) will again bind, implying that $w_0 = ar^*$. However, incentive compatibility now requires a strictly positive incentive wage, $w > 0$. Thus, w represents the agency cost per unit of throughput, and it is the behavior of w in b that captures the incentive effect of b in the Second-Best setting. Note that while the principal could infer the agent's effort from the *mean* steady-state throughput, only the *realized* throughput, $\tilde{K}$, is observed, upon which the principal cannot perfectly infer the agent's choice of effort.

Lemma 1 characterizes this incentive effect and guarantees the validity of the First-Order Approach to the agent's choice of effort:

Lemma 1: The agent's problem is concave in r. Further, the optimal incentive wage is:

$$w(b,r^*) = a\left(\frac{\partial K(b,r)}{\partial r}\right)^{-1}\Bigg|_{r=r^*}$$

which is non-negative and increasing in the principal's choice of r^*.

The optimal incentive wage is given by the agent's marginal cost of effort divided by the sensitivity of mean steady-state throughput to his effort. The more sensitive the performance metric is to the agent's effort (i.e., the greater the agent's marginal productivity), the lower the piece-rate required to induce the agent to implement the obedient effort and the lower the agency cost of doing so. The basic idea is that the more sensitive throughput is to variations in the agent's setup effort, the more likely it is that the principal can detect any deviation from the desired setup effort, r^*; therefore, the lower the incentive wage necessary to dissuade the agent from shirking. Henceforth, we refer to the sensitivity of throughput to the agent's effort, $\frac{\partial K(b,r)}{\partial r}$, as the *informa-*

[12] Later we will discuss more fully how our results would manifest in the principal's full expected profit-maximization problem.

[13] See Demougin and Garvie (1991) for a similar assumption. Alternatively, we could assume that, because of the agent's limited liability, the principal cannot use negative payments. Our results remain unchanged with this alternative formulation. The only effect is on the optimal fixed wage.

tiveness of throughput with respect to the agent's effort.[14] As illustrated by (1), the agent's ability to affect throughput is entirely driven by his ability to affect the probability of blocking and starving via his choice of effort. Given that the wage required to induce the agent to take the desired action is inversely related to the informativeness of throughput, we turn our attention to characterizing the relation between the principal's choice of the buffer size b and the informativeness of throughput with respect to the agent's action.

III. RESULTS

In what follows, we will sometimes refer to the mean rate at which the workstation processes the intermediate units relative to their mean arrival rate, $r(1 + h)/\lambda$, as ρ, the agent's *relative* effort or the *production schedule*. We parameterize our discussion with respect to the agent's relative effort since the amount of help resources, h, and the mean arrival rate of the intermediate units, λ, only affect the principal's incentive problem as they relate to the agent's effort via ρ.

As noted earlier, the principal may choose to avoid excessive blocking and starving by inducing the agent to set the workstation to process intermediate units at a rate exceeding that at which they arrive. Therefore, consider first the case in which $\rho > 1$. In the agency model, the principal compensates the agent as if she were inferring the agent's effort from the performance metric (Holmstrom 1979). But recall that throughput is subject to two sources of uncertainty that make inferring the agent's setup effort from it difficult: the stochastic arrival rate of the intermediate units and the workstation's stochastic processing rate. If the buffer size, b, is sufficiently small, then, in expectation, the probability of starving is large, resulting in relatively few units per period available for the workstation to process into throughput (i.e., a small potential sample size). Therefore, when $\rho > 1$ and the buffer is small, throughput is relatively uninformative about the agent's effort. By *increasing* the buffer, b, the principal can increase the expected number of units available to the workstation, thereby making throughput less sensitive to the stochastic arrivals of the intermediate units and more sensitive to the agent's setup effort. Thus, when b is small, the principal can improve the informativeness of throughput, and reduce the compensation weight on throughput, by increasing b.

The above intuition is based on a small increase in b starting at a buffer that, given $\rho > 1$ (i.e., $r(1+h) > \lambda$), is sufficiently small to severely restrict throughput due to blocking and starving. However, the effect of increasing the buffer on the informativeness of throughput differs once the buffer is already large enough so that blocking and starving are both unlikely. Increasing buffer size beyond some point must result in a *decrease* in the informativeness of throughput with respect to the agent's effort choice. To see this, note that in the extreme, when the buffer is infinite, there is zero probability of blocking and, given $\rho > 1$, the mean steady-state throughput is $K = \lambda$, which is *independent* of the agent's effort in excess of λ. Thus, when there is no blocking, throughput is *totally* insensitive to and uninformative about any agent effort beyond the cut-off level λ. This implies that some probability of blocking is necessary for the throughput to be informative with respect to the agent's effort. Therefore, increasing b beyond some point decreases the relative amount of information about the agent's setup effort. So when b is relatively large, the principal can increase the informativeness of throughput by *decreasing* b.

The intuition above suggests that increasing the size of the buffer may cause the informativeness of throughput to initially increase and later decrease. The following proposition confirms our intuition and guarantees that once the informativeness of throughput is decreasing in the size of the buffer, a larger buffer will never make throughput more informative.[15]

[14] Our notion of informativeness is closely related to that in Laux (2001) and Laffont and Martimort (2002, 156).
[15] All proofs are in Appendix B.

Proposition 1: For any fixed $\rho > 1$, the informativeness of throughput with respect to the agent's choice of effort, r, is either decreasing in the size of the buffer, b, or initially increasing in b and then decreasing in b.

Proposition 1 supports the preceding intuition by establishing that when buffers are relatively small (large), the informativeness of throughput is increasing (decreasing) in buffer size. For example, consider the case in which $h = 0$, $\lambda = 1$, and $r^* = 1.1$. The informativeness of throughput $\left(\frac{\partial K(b,r)}{\partial r}\Big|_{r=r^*} \right)$ is 0.226757 for $b = 1$, 0.292075 for $b = 2$, and 0.0286627 for $b = 50$. Thus, when $h = 0$, $\lambda = 1$, and $r^* = 1.1$, informativeness is initially increasing then decreasing in the buffer size. The proposition also establishes the possibility that the informativeness of throughput is monotonically decreasing in buffer size. This case arises because no matter how small the buffer size ($b \geq 1$), there always exists a sufficiently large level of setup effort that would cause the probability of blocking and starving to be very small, in which case the informativeness of throughput would always decrease in buffer size. For example, consider a case where the workstation is set up to process intermediate units twice as fast as they arrive: $r^* = 2$ and $h = 0$, $\lambda = 1$. The informativeness of throughput is 0.111111 for $b = 1$, 0.102041 for $b = 2$, and 1.088×10^{-14} for $b = 50$; i.e., the informativeness of throughput is always decreasing in buffer size.

The intuition for Proposition 1 suggests that, when a relatively large buffer causes the probability of blocking to be very small, the principal would want to decrease the buffer, while when a relatively small buffer causes the probability of blocking to be very large, the principal would want to increase the buffer. Since the probability of blocking is always decreasing in the agent's relative effort, ρ, this intuition suggests that the optimal buffer size should also be decreasing in ρ. This logic is formalized in Proposition 2:

Proposition 2: For $\rho > 1$, the buffer size, b, which maximizes the informativeness of throughput with respect to the agent's choice of effort, is decreasing in the agent's relative effort, ρ.

Lemma 1 and Proposition 1 lead to the empirically testable result that the incentive wage rate (i.e., the marginal cost of effort divided by the informativeness of throughput) should be either increasing or U-shaped in buffer size. Combining Lemma 1 with Proposition 2 implies that one would expect to see smaller buffers in settings with relatively aggressive production schedules; i.e., $\rho \gg 1$.

To this point, we have only addressed the case in which the principal wants the agent to set the workstation to operate, on average, at a faster rate than intermediate goods arrive ($\rho > 1$). When $\rho < 1$, on average, the workstation processes the intermediate units at a slower rate than they arrive. Unlike the case with $\rho > 1$, when $\rho < 1$, even as the buffer size tends to infinity, the marginal effect of the agent's effort on the probability of blocking does not go to zero, because no matter how large the buffer, the principal can never eliminate the threat of blocking. In this setting, the principal continues to increase the informativeness of throughput by increasing b, as this lowers the probability of blocking, which in turn allows the agent a greater opportunity to influence the probability of blocking and, consequently, throughput. This logic is summarized in Corollary 1:

Corollary 1: When $\rho < 1$, the informativeness of throughput, K, with respect to the agent's effort, r, is everywhere increasing in the buffer size, b.

The above corollary suggests that minimizing buffers, as advocated by the JIT literature (Schonberger 1986, 67) exacerbates the incentive problem when $\rho < 1$. However, when $\rho > 1$, minimal buffers ($b = 1$) may be optimal from an incentive perspective under certain conditions. Because buffer sizes are integer-valued, when $\rho > 1$, it is difficult to characterize the optimal buffer in closed form, though Proposition 3 establishes an interval in which the optimal b must lie.

Proposition 3: When $\rho > 1$:

a. the optimal b lies in the interval: $\max\left(\frac{1}{\sqrt{\rho-1}},2\right) - 1 \leq b \leq \max\left(\frac{1}{\rho-1},1\right)$.
b. $\rho \geq 2$ is a sufficient condition for minimal buffers ($b = 1$) to be optimal.

The first part of Proposition 3 indicates that the common objective in the JIT literature of minimizing buffers would incur unnecessary agency costs if the agent is not being asked to work relatively hard; i.e., $1 < \rho < 2$. More generally, our results point to the complementarities that are created between the choice of buffer size and the choice of the agent's relative effort, ρ, as a result of the effect of both on the informativeness of throughput as the performance metric.[16] As we have just shown, changes in buffer size without associated changes in the agent's desired relative effort can lead to unintended agency costs.

Our analysis of the effect of buffer size on the informativeness of throughput is based on the role of buffers in blocking and starving the production process, with the other organizational design variables (r^*,λ,h) held fixed. However, we can also view these design variables in terms of their potential effects on the informativeness of throughput via *their* effects on the probabilities of blocking and starving. Situations can exist in which it may be more efficient to manipulate the informativeness of throughput by varying these other design choices than by varying the buffer size. For example, increasing buffer size has space utilization implications that the other variables do not; i.e., one cannot increase buffer size if there is no additional space available, whereas it may be possible to upgrade the workstation machinery (increase h) or speed up/slow down the upstream workstations (vary λ). Similarly, buffer size increments come in discrete steps while the other design variables may allow for finer adjustments. This motivates us to consider the effect of these other design choices on the informativeness of throughput while holding the buffer size fixed, as considered next.

Proposition 4: Holding the buffer fixed, the informativeness of throughput with respect to the agent's setup effort is:

a. increasing in the mean arrival rate of intermediate units, λ,
b. decreasing in the principal's desired level of agent effort, r^*, and
c. single-peaked in the level of help resources allocated to the agent, h.

Increasing the arrival rate, λ, decreases the likelihood of starving, which in turn reduces the effect of the variation in arrivals on throughput, making throughput more informative about the other source of variation, the agent's effort, r. The same logic holds for decreases in r. The more interesting result concerns productivity-enhancing investments, h. Proposition 4 implies that even ignoring out-of-pocket costs, there are two counter-acting effects to increasing productivity-enhancing investments for the agent. The direct positive effect is that such investments make the agent more productive, thereby increasing his ability to influence throughput via blocking and starving. The indirect, negative effect is that such investments may render the agent *so* productive, that holding effort fixed, the threat of blocking tends to zero as h increases, even if the agent's effort is only negligible. As (1) indicates, throughput only reflects effort to the extent that the agent's effort affects the probability of blocking. Beyond some point, increases in h make throughput less sensitive to the agent's effort and therefore less informative about the agent's effort, making it more costly to motivate him. In a sense, resources, h, play the same role as buffer size in our earlier analysis. Both h and b can be looked at as productivity-enhancing investments and

[16] Thus, our results are consistent with the complementarities literature, for example, Milgrom and Roberts (1990, 1992).

both can be the source of additional agency costs as a result of their effects on the informativeness of throughput as a performance metric.

One can interpret our results as highlighting a value to creating bottlenecks, either by choosing a finite buffer size or a finite amount of productivity-enhancing investments. Because we have abstracted from both the out-of-pocket costs associated with inventory buffers, ($c > 0$) and the effect of the principal's choice of buffer size on the expected throughput and revenue, these results arise solely for incentive reasons. Allowing for these two additional effects will mitigate the force of our results; however, because all of our findings are obtained using arbitrary parameter values, the incentive issues identified would persist. Appendix A provides a numerical example where we solve the principal's more general, expected steady-state, profit-maximizing problem. The solution illustrates that the incentive effect of buffer size identified will continue to hold in this more general setting.

IV. CONCLUSION

The present study analyzes the relation between the choice of the firm's production line buffers and optimal incentives. The operations management literature has focused on the importance of minimizing buffers and reducing inventory. A major insight from the present work is that we identify an additional incentive value associated with the firm's choice of buffers. We show that under reasonable conditions, the informativeness of throughput as a performance metric is single-peaked in both buffer size and production-enhancing resources, such that it is optimal to induce non-zero, albeit bounded, probabilities of blocking and starving. We show further that the information-maximizing buffer is decreasing in the effort that the principal wants to induce from the agent. Therefore, we would expect to observe smaller buffer sizes in settings with more aggressive production schedules (i.e., $\rho \gg 1$). Our findings thus emphasize the incentive complementarity between buffer size and production schedules, and provide a possible explanation for the mixed empirical evidence regarding the profitability of work-in-process inventory reductions (Nagar et al. 2009).

This research spans two different disciplines, managerial accounting and operations management, each with its own nuances and paradigms. It represents an attempt to bridge these two literatures by weaving together some of the basic models of each. To do so, we made a number of modeling assumptions. We ignored some of the factors that are identified in the operations management literature as influencing the inventory buffer decision in order to focus on incentive issues and the informational effect of the buffer size decision. For example, some of the claimed benefits of small buffers under JIT that we have ignored include the increased ability to identify production problems early, shorter lead-times, and faster responsiveness to changes in market conditions. However, we are confident that the incentive consequences of inventory buffers brought to light by our analysis will be present notwithstanding these other factors.

A limitation of our model is our restricted contracting space. While we limited our analysis to linear contracts for analytical reasons, the crucial assumption underlying our results is the monotonicity of compensation in the agent's marginal expected productivity, which is either assumed or shown to be optimal in the majority of agency models. Our analysis also assumes that the agent is risk-neutral. However, recall that the principal chooses b based on the information content of realized throughput with respect to the agent's effort. Therefore, to the extent that a risk-averse agent must be compensated for risk, intuition would suggest that our results would be strengthened in the presence of risk-aversion. However, introducing risk-averse agents imposes several analytical complications, and we leave the resolution of such issues to future research.

Another simplifying assumption is that the agent's compensation is based on a single performance metric. Clearly, other performance metrics are possible, including realized inventory levels.[17] However, the mean steady-state inventory level does not offer any information above and beyond the mean steady-state throughput, K, since both measures are functions of exactly the same variables. On the other hand, if we allowed for *continuous* monitoring, as with radio frequency identification (RFID), then we could isolate the agent from the stochasticity of the arrival process and mitigate the agent's moral hazard problem.[18] In this case, our problem reduces to the traditional moral-hazard setting where the agent always has work to process, and the principal infers the agent's effort via stochastic output. It would therefore be interesting to empirically analyze the incentive effect of buffers both before and after the installation of an RFID monitoring application to see the extent to which they are used as substitutes to buffers for filtering noise in available performance measures. We leave to future research such inquiries, as compensation data surrounding RFID installations are not yet readily available.

APPENDIX A

Our results are based on analyzing a model in which the principal minimizes her cost of inducing the agent to exert a particular setup effort. In particular, our analysis of the optimal buffer size ignored both the effect of buffers on the out-of-pocket cost of maintaining inventory and, more importantly, the direct revenue effect of the choice of buffer size. In this appendix we show that the insights and qualitative results derived from our analysis continue to hold when we take into consideration both of these additional effects. To do so, we provide a numerical example in which the principal's objective is to maximize her steady-state expected profit. As before, her choice variables are buffer size, b, and the compensation contract, (w_0,w), determining the desired level of agent effort, r^*, which is subject to moral hazard. Including the explicit revenue and maintenance cost effects associated with the choice of buffer size, the principal's problem is now:

$$Max_{\{r^*,w,w_0,b\}} RK(b,r^*) - wK(b,r^*) - w_0 - cb$$

s.t.

$$w\tilde{K} + w_0 - ar^* \geq 0 \quad \forall\, \tilde{K} \quad \text{(LL)}$$

$$r^* \in \arg\max_r (wK(b,r) + w_0 - ar) \quad \text{(IC)}$$

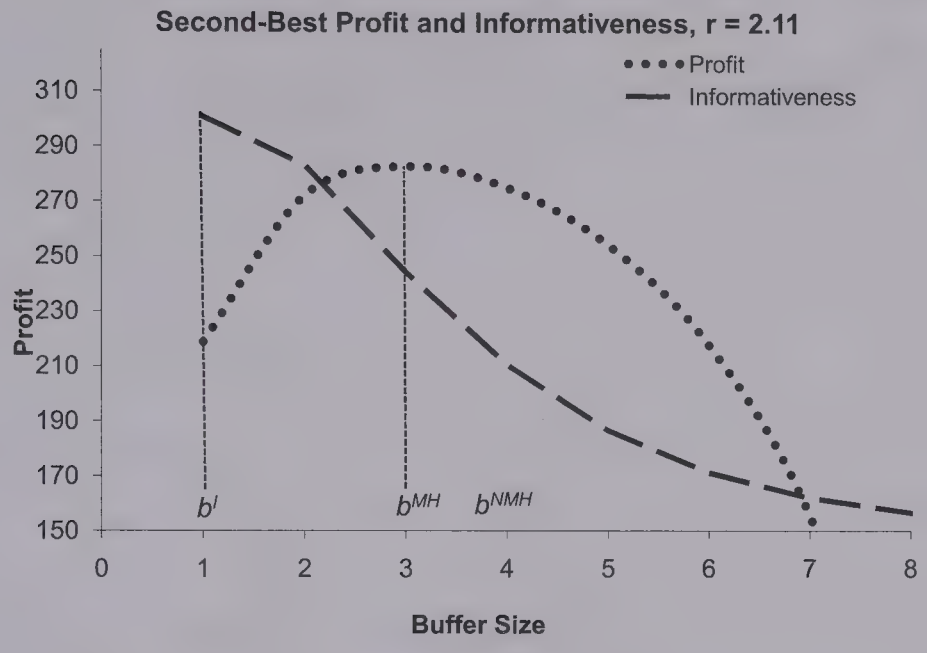
where we again assume a linear demand schedule for throughput. We first solve the above problem to find the compensation-buffer size-effort quadruple (w,w_0,b^{MH},r^*). We then solve the same problem where we again induce the agent to choose r^*, but we assume that the agent's effort is contractible and therefore not subject to moral hazard. Thus, in this second problem we are only solving for the optimal compensation and buffer size. We label the optimal buffer size in this second, no moral hazard problem, b^{NMH}. Comparing the two optimal buffer sizes allows us to isolate the incentive effects of the principal's choice of buffer size within this more general profit-maximizing setting. The example demonstrates that our findings continue to hold. In particular, the incentive effect of buffers continues to exist and can bias the principal's profit-maximizing buffer away from the no moral hazard solution.

---

[17] We thank a reviewer for raising this issue.

[18] RFID involves "an integrated circuit with an antenna, known as a 'tag,' attached to a ...product. Product information as well as other relevant information can be stored in the tag ... Using wireless technologies, readers can be set up to read the information on the tags" (Lee and Ozer 2007, 40).

FIGURE 2
Second-Best Profit and Informativeness
r = 2.11



For $R = 350$, $c = 10$, $a = 1$, $\lambda = 1$, the solution to the profit-maximizing problem with moral hazard sets the agent's effort at $r^* = \rho = 2.11$ and the optimal buffer size at $b^{MH} = 3$, though a buffer of size $b^I = 1$ maximizes the informativeness of throughput when $\rho = r^* = 2.11$. For the no moral hazard case in which the principal chooses $\rho = r^* = 2.11$, her profit (not drawn) is maximized at the buffer size $b^{NMH} = 4$.

The example assumes that the principal's revenue per unit of throughput is R dollars. In our example we set $R = 350$, $c = 10$, $a = 1$, $h = 0$ and $\lambda = 1$. In the moral hazard problem, the principal finds it optimal to induce the agent to work relatively hard ($\rho = r^* = 2.11$) and to set the buffer size at $b^{MH} = 3$.

Figure 2 plots both the expected steady-state profits for the moral hazard case and the informativeness of throughput, while fixing the induced effort at $r^* = 2.11$ and varying the buffer size. Consistent with Proposition 3 and the fact that $\rho = r^* = 2.11 > 2$, informativeness is everywhere decreasing in the size of the buffer, implying that the optimal buffer from an incentive perspective, which we label b^I, is equal to 1. However, unlike in Proposition 3, the optimal buffer size is $b^{MH} = 3$. The reason for this difference is that we are now trading off the revenue and holding cost effects of buffer size (the smaller the buffer size, the smaller the expected revenue and the lower the holding costs) against the agency cost (the smaller the buffer size, the more informative is throughput and, hence, the smaller the agency cost). The logic of Proposition 3, however, indicates that when $r^* = 2.11$, the optimal buffer size in the moral hazard case should be *smaller* than in the no moral hazard case. In fact, that is exactly what we find: the profit-maximizing buffer size in the no moral hazard case is $b^{NMH} = 4$, as indicated in Figure 2.

APPENDIX B

Lemma 0

The function $f \in C^k[a,b]$ has a root of multiplicity k at r if and only if:

$$0 = f(r) = f'(r) = f''(r) = \ldots = f^{(k-1)}(r) \quad \text{but} \quad f^{(k)}(r) \neq 0.$$

Proof of Lemma 0

See Theorem 3.20 in Vinberg (2001) and Corollary 3.21 in Artin (1991, 508). ■

Proof of Lemma 1

The agent's utility is given by $u = wK(b,r) + w_0 - ar$. Hence:

$$\frac{\partial u}{\partial r} = w\frac{\partial K}{\partial r} - a$$

$$\frac{\partial^2 u}{\partial r^2} = w\frac{\partial^2 K}{\partial r^2}.$$

Substituting the functional form of K from (1) yields:

$$\frac{\partial^2 u}{\partial r^2} = (w)\frac{\partial^2 K}{\partial r^2}$$

$$= (1+h)^2(w)(b+1)(\lambda/r(1+h))^{b+1}$$

$$\times \frac{(b(r(1+h)-\lambda)(\lambda(\lambda/r(1+h))^b + r(1+h)) + 2r(1+h)\lambda((\lambda/r(1+h))^b - 1))}{(\lambda(\lambda/r(1+h))^b - r(1+h))^3},$$

and the denominator is non-positive if and only if $r(1 + h) \geq \lambda$. Further, $(w)(b+1)(\lambda / r(1+h))^{b+1} > 0$. Hence, if we can show that $(b(r(1+h)-\lambda)(\lambda(\lambda / r(1+h))^b + r(1+h)) + 2r(1+h)\lambda((\lambda / r(1+h))^b - 1))$ is non-negative if and only if $r(1 + h) \geq \lambda$, we will then have established that $\frac{\partial^2 u}{\partial r^2} \leq 0$.

To this end, we can re-write the expression as:

$$(r(1+h))^{-b}[b(r(1+h))^{2+b} + (r(1+h))^{b+1}(-b\lambda - 2\lambda) + r(1+h)\lambda^{b+1}(2+b) - b\lambda^{b+2}]$$

labeling the bracketed expression, N. According to Descartes' rule, N can have at most three roots in r. Using Lemma 0, note that:

$$N|_{r=\frac{\lambda}{1+h}} = \left.\frac{\partial N}{\partial r}\right|_{r=\frac{\lambda}{1+h}} = \left.\frac{\partial^2 N}{\partial r^2}\right|_{r=\frac{\lambda}{1+h}} = 0 < b(1+h)^3\lambda^{b-1}(2+3b+b^2) = \left.\frac{\partial^3 N}{\partial r^3}\right|_{r=\frac{\lambda}{1+h}},$$

which shows that all three roots are at $r = \lambda/(1 + h)$. Finally, since the third derivative is positive, the Extremum Test assures that N has a saddle point at $r = \lambda/(1 + h)$;[19] therefore N alternates in sign as r crosses $\lambda/(1 + h)$. However, since N is negative at $r(1 + h) = 0$, N must be non-negative for $r(1 + h) \geq \lambda > 0$.

The constraint (IC) can be expressed as:

$$\frac{\partial}{\partial r}\left(w\frac{1-(r(1+h)/\lambda)^{-b}}{1-(r(1+h)/\lambda)^{-(b+1)}}\lambda - ar\right) = 0,$$

implying that:

[19] For an explanation of the Extremum Test, see http://mathworld.wolfram.com/ExtremumTest.html.

$$w(b,r) = a\left(\frac{\partial K(b,r)}{\partial r}\right)^{-1} = \frac{a((r(1+h)/\lambda)^{b+1}-1)^2}{(1+h)(1-(r(1+h)/\lambda)^b(1-b(1-r(1+h)/\lambda)))},$$

which is both positive and increasing in r. ■

Proof of Proposition 1

Because we have defined informativeness as a constant times the inverse of the incentive wage, it suffices to show that the incentive wage is either U-shaped in buffer size, or increasing in b. In particular, we will show that if the wage is increasing in b, at say $\hat{b}$, then the wage continues to increase for all $b > \hat{b}$. Then we will show that there exists a finite b beyond which the wage must be increasing in b.

In general, we can write:

$$w(b,\rho) = \frac{a(\rho^{b+1}-1)^2}{(1+h)(b\rho^{b+1}-(b+1)\rho^b-1)}, \tag{2}$$

where, henceforth to simplify, we omit the constant $\frac{a}{(1+h)}$. Hence:

$$\frac{\partial}{\partial b}w(b,\rho) = \rho^b(\rho^{b+1}-1)\frac{-(\rho-1)(\rho^{b+1}-1)+(b(\rho-1)(1+\rho^{b+1})-\rho(\rho^b-2)-1)\log(\rho)}{(b\rho^{b+1}-(b+1)\rho^b+1)^2}. \tag{3}$$

We first claim that the denominator of (3) is strictly positive. To see this, note that $b\rho^{b+1}-(b+1)\rho^b+1$ has a root at $\rho = 1$; its first derivative evaluated at $\rho = 1$ is zero and its second derivative is strictly positive for all $\rho \geq 1$, implying that the denominator is strictly positive for all $\rho > 1$. Thus, for $\rho > 1$, the sign of $\frac{\partial}{\partial b}w(b,\rho)$ is determined by the function $t(b,\rho)$, where:

$$t(b,\rho) = -(\rho-1)(\rho^{b+1}-1)+(b(\rho-1)(1+\rho^{b+1})-\rho(\rho^b-2)-1)\log(\rho).$$

We will show that if $t(b,\rho)$ is positive, it is increasing in b, implying that once the wage is increasing in buffer size b, it will continue to increase in b for larger buffers. To this end, note that the first term in $t(b,\rho)$ is strictly negative for $\rho > 1$. We claim that the second term in $t(b,\rho)$ is strictly positive for $\rho > 1$. To see this, note that we can write the coefficient of $\log(\rho)$ as $-1-b+(2+b)\rho-(b+1)\rho^{1+b}+b\rho^{2+b}$, which at $\rho = 1$, is both equal to 0 and increasing in ρ. Moreover, the second derivative in ρ is strictly positive for $\rho > 1$, since $b(1+b)(\rho^{b-1}(2\rho+b(\rho-1)-1) > 0$.

Thus, if $t(b,\rho)$ is positive, it must be the case that:

$$\log(\rho) > \frac{(\rho-1)(\rho^{b+1}-1)}{b(\rho-1)(1+\rho^{b+1})-\rho(\rho^b-2)-1}. \tag{4}$$

To show that if $t(b,\rho) > 0$, $t(b,\rho)$ is increasing in b, we difference $t(b,\rho)$ in b:

$$t(b+1,\rho)-t(b,\rho) = (\rho-1)(-(\rho-1)\rho^{1+b}+(1-(b+1)\rho^{1+b}+(b+1)\rho^{2+b})\log(\rho)). \tag{5}$$

The coefficient of $\log(\rho)$ in (5) is always positive, hence we can use the bound for $\log(\rho)$ from (4) to form a lower bound for $t(b+1,\rho) - t(b,\rho)$:

$$t(b+1,\rho)-t(b,\rho) = (\rho-1)(-(\rho-1)\rho^{1+b}+(1-(b+1)\rho^{1+b}+(b+1)\rho^{2+b})\log(\rho))$$

$$> (\rho-1)\left(-(\rho-1)\rho^{1+b}+(1-(b+1)\rho^{1+b}+(b+1)\rho^{2+b})\frac{(\rho-1)(\rho^{b+1}-1)}{b(\rho-1)(1+\rho^{b+1})-\rho(\rho^b-2)-1}\right)$$

$$= (\rho-1)^2\frac{-1+(3+2b)\rho^{1+b}-(3+2b)\rho^{2+b}+\rho^{3+2b}}{-1-b+\rho(2+b)-(b+1)\rho^{1+b}+b\rho^{2+b}}.$$

The denominator above is identical to that in (4), which we have already shown to be strictly positive for $\rho > 1$. Thus, we need only sign the numerator, which has at most three roots according to Descartes' rule. However, in accordance with Lemma 0, the numerator has all three roots at $\rho = 1$, and since it has a strictly positive third derivative in ρ, the expression is strictly positive for $\rho > 1$, implying that if at some $b = \hat{b}$ the wage is increasing in buffer size ($t(\hat{b},\rho) > 0$), then it will continue to be increasing in buffer size for larger buffers.

To prove that the wage is eventually increasing in buffer size, note that for an arbitrary, finite $\rho > 1$, $\lim_{b\to\infty} t(b,\rho) = \infty$, implying that there exists a finite $\tilde{b}$ for which $t(\tilde{b},\rho) > 0$. ■

Proof of Proposition 2

In order to characterize the most informative buffer size, it suffices to characterize the lowest incentive compatible wage, w. We will show that if the wage is minimized at $\hat{b}$ when $\rho = \hat{\rho}$, then the wage will be increasing in b at $\hat{b}$ when $\rho > \hat{\rho}$; thus by Proposition 1, the wage-minimizing buffer for $\rho > \hat{\rho}$ is less than or equal to $\hat{b}$.

Suppose the wage is minimized at $\hat{b}$ when $\rho = \hat{\rho}$, Then, by definition:

$$w(\hat{b},\hat{\rho}) - w(\hat{b}-1,\hat{\rho}) \le 0 \le w(\hat{b}+1,\hat{\rho}) - w(\hat{b},\hat{\rho}).$$

To complete the proof, we will prove that $w(\hat{b}+1,\hat{\rho}) - w(\hat{b},\hat{\rho}) \ge 0$ implies $w(\hat{b}+1,\rho) - w(\hat{b},\rho) > 0$ for $\rho > \hat{\rho}$.

To this end, consider the difference $w(b + 1,\rho) - w(b,\rho)$:

$$(\rho-1)\rho^b \frac{1+b-3\rho(b+1)+3\rho^{2+b}+\rho^{3+b}-(2+b)\rho^{3+2b}+b\rho^{4+2b}}{(1-(b+1)\rho^b+b\rho^{1+b})(1-(2+b)\rho^{1+b}+(b+1)\rho^{2+b})}. \quad (6)$$

We first show that the denominator of (6) is always positive. Using Descartes' rule, the first term has at most two roots. The term and its first derivative in ρ evaluated at $\rho = 1$ are both equal to 0. Therefore, Lemma 0 implies that its only positive root is $\rho = 1$ and since its leading coefficient is positive, the term is strictly positive for $\rho > 1$. The same argument shows that the second term in the denominator is strictly positive for $\rho > 1$.

Thus, the sign of (6) is determined by the numerator. Applying Descartes' rule, the numerator has at most four positive roots in ρ. Because the numerator and both its first and second derivatives in ρ are equal to 0 when $\rho = 1$, Lemma 0 implies that $\rho = 1$ is a root with multiplicity three. Because the third derivative of the numerator in ρ evaluated at $\rho = 1$ is negative $(-3b^2 - 9b - 6)$ but the leading coefficient is positive, the final root in ρ, which we label $\rho(b)$, must be greater than 1.

Now, suppose $w(\hat{b}+1,\hat{\rho}) - w(\hat{b},\hat{\rho}) \ge 0$. Then it must be the case that $\hat{\rho}$ is greater than or equal to the final root, $\rho(\hat{b})$; therefore, if $\rho > \hat{\rho}$, then $\rho \ge \rho(\hat{b})$, in which case $w(\hat{b}+1,\rho) - w(\hat{b},\rho) > 0$, and thus the wage-minimizing buffer for $\rho > \hat{\rho}$ must be no larger than $\hat{b}$. ■

Proof of Corollary 1

The proof is similar to that of Proposition 2 and is omitted. ■

Proof of Proposition 3

By definition, the information-maximizing b will equivalently minimize the wage, w. We identify a sufficient upper bound on b, as parameterized by ρ, for $w(b + 1,\rho) - w(b,\rho)$ to be positive, and a sufficient lower bound on b for $w(b + 1,\rho) - w(b,\rho)$ to be negative. Combining the

two bounds, we identify a region for b where $w(b+1,\rho) - w(b,\rho)$ crosses zero in b, defining the wage-minimizing buffer size.

We begin with two additional Lemmas.

Lemma 2

Let $f(x) = a_n x^n + \ldots + a_1 x + a_0$ be a real polynomial with $f(1) = 0$. Furthermore, assume that there exists a $k < n - 1$ such that for $i > k$, $a_i > 0$ and for $i \leq k$, $a_i \leq 0$. Then $f(x) > 0$ for $x > 1$ and $f(x) < 0$ for $x < 1$.

Proof of Lemma 2

For all $x > 1$ we have $f(x) > (a_n + a_{n-1} + \ldots + a_{k+1})x^{k+1} - (-a_k - a_{k-1} - \ldots - a_0)x^k$. Since $f(1) = 0$, the sum of the terms in the parentheses are equal, and we denote the sum (either in the first or the second set of parentheses) by s. Clearly, $s > 0$ and $sx^k(x-1) > 0$ for $x > 1$; hence, $f(x) > 0$ when $x > 1$. Similarly, for all $0 \leq x < 1$, we have $f(x) < sx^{k+1} - sx^k = sx^k(x-1) < 0$. This concludes the proof of the lemma. ■

Lemma 3

Let b be a natural number and $\rho > 1$, then:

$$\frac{2}{b}\rho + 1 \leq \frac{\sum_{i=0}^{b} (i+1)\rho^i}{\sum_{i=0}^{b-1} (i+1)\rho^i} \leq \frac{b+1}{b}\rho + \frac{1}{b}$$

for $\rho > 1$, and the opposite weak inequalities hold when $\rho < 1$.

Proof of Lemma 3

First consider the inequality on the right-hand side. Set:

$$f(\rho) = \left(\frac{b+1}{b}\rho + \frac{1}{b}\right)\sum_{i=0}^{b-1} (i+1)\rho^i - \sum_{i=0}^{b} (i+1)\rho^i.$$

We want to show that $f(\rho) > 0$ for $\rho > 1$ and $f(\rho) < 0$ for $\rho < 1$. Thus, it is sufficient to show that $f(\rho)$ satisfies the hypothesis of the previous lemma. First, to show that $f(1) = 0$, we have:

$$f(1) = \left(\frac{b+1}{b} + \frac{1}{b}\right)\sum_{i=0}^{b-1} (i+1) - \sum_{i=0}^{b} (i+1) = \left(\frac{b+2}{b}\right)\left(\frac{b}{2}\right)(b+1) - \frac{(b+1)(b+2)}{2} = 0.$$

We must now find a k as in the prior lemma. If $f(\rho) = a_n\rho^n + \ldots a_1\rho + a_0$, then $n = b - 1$ (although there is a ρ^b term in the construction of $f(\rho)$, note that it has a coefficient of zero) and for arbitrary $m \geq 1$, the coefficient a_m is given by:

$$\frac{b+1}{b}m + \frac{1}{b}(m+1) - (m+1) = \frac{2m - (b-1)}{b}.$$

Hence, $a_m > 0$ if and only if $m > \frac{b-1}{2}$. In particular, a threshold k exists when $n = b - 1 > \frac{b-1}{2}$, which is valid for $b \geq 2$. If $b = 1$, then $f(\rho) = \frac{1}{2}(\rho - 1)$, which is trivially greater than zero for $\rho > 1$ and less than zero for $\rho < 1$. Hence, Lemma 2 holds and we have established Lemma 3 for the right-hand side inequality.

Next, consider the left-hand side inequality in the statement of the lemma. Let:

$$g(\rho) = \sum_{i=0}^{b} (i+1)\rho^i - \left(\frac{2}{b}\rho + 1\right)\sum_{i=0}^{b-1} (i+1)\rho^i.$$

Note that:

$$g(1) = \sum_{i=0}^{b} (i+1) - \left(\frac{2}{b} + 1\right)\sum_{i=0}^{b-1} (i+1) = 0.$$

Moreover, the leading coefficient of $g(\rho)$ is positive for $b > 1$ and all the remaining coefficients are always negative; hence, the conditions of Lemma 2 are always satisfied for $b > 1$. If $b = 1$, then $g(\rho) = 2(\rho - 1)$, which is again positive for $\rho > 1$ and negative for $\rho < 1$, as was to be shown. This concludes the proof of the lemma. ■

Continuing with the proof of Proposition 3, when $\rho > 1$, we can rewrite $w(b,\rho)$ without the constant multiplier $\frac{a}{(1+h)}$ as:

$$w(b,\rho) = \frac{(\rho^{b+1} - 1)^2}{(b\rho^{b+1} - (b+1)\rho^b + 1)} = \frac{\rho^{2b} + 2\rho^{2b-1} + \ldots + (b+1)\rho^b + b\rho^{b-1} + \ldots + 2\rho + 1}{b\rho^{b-1} + (b-1)\rho^{b-2} + (b-2)\rho^{b-3} + \ldots + 2\rho + 1}.$$

To see this, we begin by rewriting the numerator:

$$\begin{aligned}
(\rho^{b+1} - 1)^2 &= ((\rho^{b+1} + \rho^b + \rho^{b-1} + \ldots + \rho) - (\rho^b + \rho^{b-1} + \ldots + 1))^2 \\
&= ((\rho - 1)(\rho^b + \rho^{b-1} + \ldots + 1))^2 \\
&= (\rho - 1)^2(\rho^b + \rho^{b-1} + \ldots + 1)^2 \\
&= (p - 1)^2\left(\sum_{i=0}^{b} \rho^i\right)^2 \\
&= \rho^{2b} + 2\rho^{2b-1} + \ldots + (b+1)\rho^b + b\rho^{b-1} + \ldots + 2\rho + 1.
\end{aligned}$$

The denominator can be rewritten as:

$$\begin{aligned}
b\rho^{b+1} - (b+1)\rho^b + 1 &= b(\rho^{b+1} - \rho^b) - (\rho^b - 1) \\
&= b\rho^b(\rho - 1) - (\rho - 1)\sum_{i=1}^{b} \rho^{b-i} \\
&= (\rho - 1)\left(b\rho^b - \sum_{i=1}^{b} \rho^{b-i}\right) \\
&= (\rho - 1)\sum_{i=1}^{b} (\rho^b - \rho^{b-i}) \\
&= (\rho - 1)^2\sum_{i=0}^{b-1} (i+1)\rho^i \\
&= b\rho^{b-1} + (b-1)\rho^{b-2} + (b-2)\rho^{b-3} + \ldots + 2\rho + 1.
\end{aligned}$$

In light of this representation, we can continue simplifying to obtain:

$$w(b,\rho) = \frac{\rho^{2b} + 2\rho^{2b-1} + \ldots + (b+1)\rho^{b} + b\rho^{b-1} + \ldots + 2\rho + 1}{b\rho^{b-1} + (b-1)\rho^{b-2} + (b-2)\rho^{b-3} + \ldots + 2\rho + 1}$$

$$= \frac{\rho^{2b} + 2\rho^{2b-1} + \ldots + (b+1)\rho^{b}}{b\rho^{b-1} + (b-1)\rho^{b-2} + (b-2)\rho^{b-3} + \ldots + 2\rho + 1} + 1$$

$$= \rho^{2b}\frac{1 + 2\rho^{-1} + \ldots + (b+1)\rho^{-b}}{b\rho^{b-1} + (b-1)\rho^{b-2} + (b-2)\rho^{b-3} + \ldots + 2\rho + 1} + 1. \tag{7}$$

Using (7), $w(b+1,\rho) > w(b,\rho)$ is equivalent to:

$$\rho^{2b+2}\frac{1 + 2\rho^{-1} + \ldots + (b+2)\rho^{-b-1}}{(b+1)\rho^{b} + \ldots + 2\rho + 1} + 1 > \rho^{2b}\frac{1 + 2\rho^{-1} + \ldots + (b+1)\rho^{-b}}{b\rho^{b-1} + \ldots + 2\rho + 1} + 1$$

$$\Leftrightarrow \rho^{2}\frac{1 + 2\rho^{-1} + \ldots + (b+2)\rho^{-b-1}}{(b+1)\rho^{b} + \ldots + 2\rho + 1} > \frac{1 + 2\rho^{-1} + \ldots + (b+1)\rho^{-b}}{b\rho^{b-1} + \ldots + 2\rho + 1}$$

$$\Leftrightarrow \rho^{2}\left(\frac{(b+1)\rho^{b} + \ldots + 2\rho + 1}{1 + 2\rho^{-1} + \ldots + (b+1)\rho^{-b}}\right)\left(\frac{1 + 2\rho^{-1} + \ldots + (b+2)\rho^{-b-1}}{(b+1)\rho^{b} + \ldots + 2\rho + 1}\right)$$

$$> \left(\frac{(b+1)\rho^{b} + \ldots + 2\rho + 1}{1 + 2\rho^{-1} + \ldots + (b+1)\rho^{-b}}\right)\left(\frac{1 + 2\rho^{-1} + \ldots + (b+1)\rho^{-b}}{b\rho^{b-1} + \ldots + 2\rho + 1}\right)$$

$$\Leftrightarrow \rho^{2}\frac{1 + 2\rho^{-1} + \ldots + (b+2)\rho^{-b-1}}{1 + 2\rho^{-1} + \ldots + (b+1)\rho^{-b}} = \rho^{2}\frac{\sum_{i=0}^{b+1}(i+1)\rho^{-i}}{\sum_{i=0}^{b}(i+1)\rho^{-i}} > \frac{(b+1)\rho^{b} + \ldots + 2\rho + 1}{b\rho^{b-1} + \ldots + 2\rho + 1} = \frac{\sum_{i=0}^{b}(i+1)\rho^{i}}{\sum_{i=0}^{b-1}(i+1)\rho^{i}}. \tag{8}$$

We can use Lemma 3 with $\rho > 1$ on the right-hand side of (8) to conclude that it is less than or equal to $\frac{b+1}{b}\rho + \frac{1}{b}$. Since we have assumed that $\rho > 1$, $\rho^{-1} < 1$, and we can again apply Lemma 3, replacing ρ with ρ^{-1} on the left-hand side of (8), to conclude that it is greater than or equal to $\rho^{2}\left(\frac{b+2}{b+1}\rho^{-1} + \frac{1}{b+1}\right)$. Thus, a sufficient condition on b and ρ for $w(b+1,\rho) > w(b,\rho)$ is given by:

$$\frac{1}{b+1}\rho^{2} + \frac{b+2}{b+1}\rho > \frac{b+1}{b}\rho + \frac{1}{b} \Leftrightarrow \frac{1}{b+1}\rho^{2} + \frac{b+2}{b+1}\rho - \frac{b+1}{b}\rho - \frac{1}{b} > 0.$$

Solving the second quadratic above for equality gives two roots in ρ: -1 and $\frac{b+1}{b}$. Since the leading coefficient is positive, we know that (8) holds whenever $\rho > \frac{1+b}{b}$, or a sufficient condition for the wage to be increasing in b is given by $b > \frac{1}{\rho - 1}$. Because our analysis applies to all $\rho > 1$, (8) implies that for $\rho \geq 2$, the wage-minimizing buffer is given by $b = 1$, the smallest feasible buffer size. This establishes part (b).

Next we establish a sufficient upper bound for the wage to be decreasing in b. When $a > 0$ and $\rho > 1$, $\rho^{-a} < 1$; thus, Lemma 3 implies that the left-hand side of (8) is less than $\rho^{2}\left(\frac{2}{b+1}\rho^{-1} + 1\right)$, whereas the right-hand side is bounded below by $\frac{2}{b}\rho + 1$. Hence, a sufficient condition for (8) to be violated (for the wage to be decreasing in b) is given by:

$$\frac{2}{b+1}\rho+\rho^2-\frac{2}{b}\rho-1<0. \tag{9}$$

The two roots to the polynomial on the left-hand side of (9) are given by $\rho=\frac{\frac{2}{b(b+1)}-\sqrt{\frac{4}{b^2(b+1)^2}+4}}{2}$ and $\rho=\frac{\frac{2}{b(b+1)}+\sqrt{\frac{4}{b^2(b+1)^2}+4}}{2}$, since the latter is negative and the polynomial has a positive leading coefficient; a sufficient condition for the wage to be decreasing in b is given by $1<\rho<\frac{\frac{2}{b(b+1)}+\sqrt{\frac{4}{b^2(b+1)^2}+4}}{2}$. A lower bound for $\frac{\frac{2}{b(b+1)}+\sqrt{\frac{4}{b^2(b+1)^2}+4}}{2}$ is given by $\frac{\frac{2}{b(b+1)}+2}{2}=1+\frac{1}{b(b+1)}$, and since $b(b+1)<(b+1)^2$, a sufficient condition for $w(b,\rho)$ to be decreasing is thus $\rho<1+\frac{1}{(b+1)^2}$ or $\frac{1}{\sqrt{\rho-1}}-1>b$. The lower bound on b is redundant once $\rho>\frac{5}{4}$, as b is bounded below by 1. ■

Proof of Proposition 4

In order to prove that the sensitivity of K to r is increasing in λ, decreasing in r, and single-peaked in h, we will show that the wage w is decreasing in λ, increasing in r and U-shaped in h. Differentiating w with respect to h yields:

$$\frac{\partial w}{\partial h}=a\frac{(-1+\rho^b(1+2b+b^2)-\rho^{b+1}(4b+b^2)+\rho^{2b+2}(1+4b)-\rho^{3b+2}(1+b)^2+\rho^{3b+3}b^2)}{(\rho^2(1+\rho^b(\rho b-b-1)))^2}. \tag{10}$$

The denominator is unambiguously non-negative, whereas Descartes' rule implies the numerator has at most five positive roots. Because the numerator and the first three derivatives with respect to ρ have a root at $\rho=1$, only one root remains unidentified. If $b=1$, then the fourth derivative of the numerator in (10) is zero at $\rho=1$, and thus the fifth root is in fact $\rho=1$. If $b>1$, then the fourth derivative of the numerator in (10) is strictly positive at $\rho=1$, hence, by the Extremum Test, $\rho=1$ is a relative minimum and there exists a sufficiently small $\varepsilon>0$ such that the numerator is positive at $\rho=1-\varepsilon$ when $b>1$. However, because the numerator in (10) is strictly negative at $\rho=0$, the final root must belong to the interval (0,1).

Differentiating w with respect to r yields:

$$\frac{\partial w}{\partial r}=-a\frac{(1+b)\rho^{b-1}(1-\rho^{b+1})(-b+\rho(b+2)-\rho^{b+1}(b+2)+\rho^{b+2}b)}{\lambda(1-\rho^b(1+b)+b\rho^{b+1})^2}. \tag{11}$$

The denominator of (11) is non-negative. Descartes' rule implies that in addition to $\rho=1$, (11) may have up to three additional positive roots. Using Lemma 0, the expression $-b+\rho(b+2)-\rho^{b+1}(b+2)+\rho^{b+2}b$ and its first two derivatives have a root at $\rho=1$, whereas the third derivative is strictly positive at $\rho=1$, implying that the expression has a saddle point at $\rho=1$. Since the pre-multiplier, $1-\rho^{b+1}$, also alternates in sign at $\rho=1$, the entire expression, $\frac{\partial w}{\partial r}$, is non-positive. Finally, differentiating w with respect to λ yields $\frac{\partial w}{\partial \lambda}=-\frac{r}{\lambda}\frac{\partial w}{\partial r}$, and since $\frac{\partial w}{\partial r}$ is non-positive, $\frac{\partial w}{\partial \lambda}$ must be non-negative. ■

REFERENCES

Alles, M., S. Datar, and R. A. Lambert. 1995. Moral hazard and management control in just-in-time settings. *Journal of Accounting Research* 33 (Supplement): 177–204.

Artin, M. 1991. *Algebra*. Upper Saddle River, NJ: Prentice Hall.

Balachandran, K., and S. Radhakrishnan. 1996. Cost of congestion, operational efficiency and management accounting. *European Journal of Operational Research* 89 (2): 237–245.

Balakrishnan, R., N. Nagarajan, and K. Sivaramakrishnan. 1998. The effect of property rights and audit information quality on team incentives for inventory reduction. *Management Science* 44 (9): 1193–1204.

Balsamo, S., V. Persone, and R. Onvural. 2001. *Analysis of Queuing Networks with Blocking*. Boston, MA: Kluwer Academic Publishers.

Berg, N., and N. Fast. 1975. The Lincoln Electric Company. Harvard Business School Case 9-376-028. Cambridge, MA: Harvard Business School.

Berkley, B. 1992. A review of the kanban production control research literature. *Production and Operations Management* (4): 393–411.

Cremer, J. 1995. Towards an economic theory of incentives in just-in-time manufacturing. *European Economic Review* 39 (3–4): 432–439.

Datar, S., and M. Rajan. 1995. Optimal incentive schemes in bottleneck-constrained production environments. *Journal of Accounting Research* 33 (Spring): 33–58.

Demougin, D., and D. Garvie. 1991. Contractual design with correlated information under limited liability. *RAND Journal of Economics* 22 (4): 477–489.

Feltham, G., and J. Xie. 1994. Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review* 69 (3): 429–454.

Gietzmann, M., and T. Hemmer. 2002. On the relation between optimal incentive structures and the cost and benefits of bottlenecks. *Journal of Labor Economics* 20 (2, part 2): 34–56.

Groenevelt, H. 1993. The just-in-time system. In *Handbooks in OR & MS*, edited by Graves, S., A. Rinnooy Kan, and P. Zipkin, 629–670. Amsterdam, The Netherlands: Elsevier.

Grossman, S., and O. Hart. 1983. An analysis of the principal agent problem. *Econometrica* 51 (1): 7–45.

Hall, B., C. Madigan, and E. Lazear. 2000. Performance pay at Safelite Auto Glass. Harvard Business School Case 800291. Cambridge, MA: Harvard Business School.

Hemmer, T. 1995. On the interrelation between production technology, job design, and incentives. *Journal of Accounting and Economics* 19 (2–3): 209–245.

———. 1998. Performance measurement systems, incentives, and the optimal allocation of responsibilities. *Journal of Accounting and Economics* 25 (3): 321–347.

Holmstrom, B. 1979. Moral hazard and observability. *The Bell Journal of Economics* 10 (1): 74–91.

———, and P. Milgrom. 1991. Multi-task principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law Economics and Organization* 7 (Spring): 24–52.

Hopp, W., and M. Spearman. 2001. *Factory Physics*. 1st edition. Chicago, IL: Irwin.

Horngren, C., S. Datar, M. Rajan, and C. Ittner. 2008. *Cost Accounting: A Managerial Emphasis*. 13th edition. Upper Saddle River, NJ: Pearson Prentice Hall.

Laffont, J.-J., and D. Martimort. 2002. *The Theory of Incentives: The Principal-Agent Model*. Princeton, NJ: The Princeton University Press.

Laux, C. 2001. Limited-liability and incentive contracting with multiple projects. *RAND Journal of Economics* 32 (3): 514–526.

Lee, H., and O. Ozer. 2007. Unlocking the value of RFID. *Production and Operations Management* 16 (1): 40–64.

Lu, L., J. Van Mieghem, and R. Savaskan. 2009. Incentives for quality through endogenous routing. *Manufacturing & Service Operations Management* 11 (2): 254–273.

Melumad, N., D. Mookherjee, and S. Reichelstein. 1995. Hierarchical decentralization of incentive contracts. *RAND Journal of Economics* 26 (4): 654–672.

Milgrom, P., and J. Roberts. 1990. The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review* 80 (3): 511–528.

———, and ———. 1992. *Economics, Organization & Management*. Englewood Cliffs, NJ: Prentice Hall.

———, and ———. 1995. Complementarities and fit: Strategy, structure, and organizational change in manufacturing. *Journal of Accounting and Economics* 19 (2–3): 179–208.

Nagar, V., M. Rajan, and R. Saouma. 2009. The incentive value of inventory and cross-training in modern manufacturing. *Journal of Accounting Research* 47 (4): 991–1026.

Radhakrishnan, S., and K. Balachandran. 1995. Delay cost and incentive schemes for multiple users. *Management Science* 41 (4): 646–652.

———, and ———. 2004. Service capacity and incentive compatible cost allocation for reporting usage forecasts. *European Journal of Operational Research* 157 (1): 180–195.

Riordan, M., and D. Sappington. 1987. Information, incentives, and organizational mode. *The Quarterly Journal of Economics* 2 (May): 243–264.

Schonberger, R. 1986. *World Class Manufacturing: The Lessons of Simplicity Applied*. New York, NY: The Free Press.

Vinberg, E. B. 2001. *A Course in Algebra*. Vol. 56. Providence, RI: American Mathematical Society.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 1861–1886

American Accounting Association
DOI: 10.2308/accr.2010.85.6.1861

Supervisor Discretion in Target Setting: An Empirical Investigation

Jasmijn C. Bol
University of Illinois at Urbana–Champaign

Timothy M. Keune
University of South Carolina

Ella Mae Matsumura
University of Wisconsin–Madison

Jae Yong Shin
Seoul National University

ABSTRACT: In a setting in which corporate headquarters dictates total sales targets, we study how supervisors allocate sales targets to individual stores. Specifically, we analyze whether supervisors strategically use discretion in the target-setting process to address compensation contracting issues. We first examine whether supervisors use discretion to manage compensation risk. The results are consistent with the agency-theoretic prediction that supervisors provide easier targets to stores facing higher levels of store-specific risk. Next, we examine whether discretion is used to mitigate fairness concerns. The results suggest that, consistent with behavioral arguments, supervisors use discretion to deal with fairness issues, even if the area of the supervisor's discretion is not the source of the fairness concerns. Finally, we analyze whether supervisors use discretion in the target-setting process to reduce their potential confrontation costs. Consistent with research in psychology, we find that supervisors provide easier targets to store managers with relatively higher hierarchical status.

Keywords: *discretion; sales target; target-setting process; compensation risk; fairness concerns; confrontation costs.*

We are grateful to the participating company for allowing us access to the data. We are especially grateful to Joong-Bum Choi for providing us with institutional insights and helpful suggestions. We also thank Shannon Anderson (editor), Steve Kachelmeier (senior editor), two anonymous reviewers, Marsha Keune, Joan Luft, David Marginson, Brian Mayhew, Frank Moers, Chad Proell, Hanno Roberts, Lloyd Tanlu, Michael Williamson, Dae-Hee Yoon, and participants at the 6th EIASM Conference on New Directions in Management Accounting, 2009 AAA Management Accounting Section Conference, 2009 Global Management Accounting Research Symposium, 2009 AAA Annual Meeting, and workshops at Georgia State University, IESE Business School, the University of Illinois at Urbana–Champaign, University of Wisconsin–Madison, and VU University Amsterdam for their helpful comments. We also thank Ho-Sang Shin for his able research assistance. Professors Bol and Shin gratefully acknowledge research support from the Department of Accountancy at the University of Illinois at Urbana–Champaign, and Professors Keune and Matsumura gratefully acknowledge support from the Department of Accounting and Information Systems at the University of Wisconsin–Madison.

Editor's note: Accepted by Shannon Anderson.

Submitted: March 2009
Accepted: April 2010
Published Online: November 2010

Data Availability: *Data in this study are derived from a proprietary source.*

I. INTRODUCTION

This study examines supervisor discretion in target setting. Specifically, in a context in which overall regional sales targets are dictated by corporate headquarters, we analyze how regional directors (supervisors) use discretion to allocate regional sales targets to individual stores in the region. Understanding how performance targets are set is important because such targets play a key role in many aspects of management accounting and control. For instance, in the area of performance evaluation, targets communicate supervisors' expectations and directly influence employee incentives, as meeting or exceeding targets leads to increased short- and/or long-term compensation. Despite the importance of targets, however, our knowledge of supervisor target-setting behavior is limited.

A considerable amount of research examines the related topic of goal setting. Much of this literature addresses the relationship between target achievability and employee effort or performance (see, e.g., Locke and Latham 1990). The results from this goal-setting literature, however, do not provide clear insight into supervisors' target-setting behavior. For example, Merchant and Manzoni (1989) find that supervisors set targets at substantially more attainable levels (i.e., at levels with a higher probability of achievement) than the goal-setting theory predicts. The authors' interview data indicate that the targets are more attainable than predicted because eliciting high levels of employee performance is not supervisors' only concern. Other factors, such as increasing the predictability of budgets or discouraging earnings management, also influence supervisors' target-setting decisions. To date, however, there exists little empirical evidence on the array of factors influencing supervisors' target-setting decisions and methods (Ittner and Larcker 2001; Anderson et al. 2010). We address this gap in the literature by analyzing how directors strategically use discretion in the target-setting process to deal with compensation contracting issues.

Specifically, we begin by examining whether directors use discretion to manage compensation risk. Consistent with predictions based on agency theory, we hypothesize that directors provide more attainable targets to stores that face higher levels of store-specific risk. Next, we study whether discretion is used to mitigate fairness concerns. Drawing on behavioral theory, we predict that directors will provide more attainable targets to stores when other elements of the compensation plan are taken to be unfair. Finally, we analyze whether directors use discretion in the target-setting process to minimize confrontation costs. Consistent with research in psychology, we hypothesize that directors provide easier targets to store managers with relatively higher hierarchical status.

We test our predictions using archival data across multiple years and organizational business units. In particular, we employ a rich dataset from the population of 103 business units (referred to as stores) managed by a postal service provider in Korea for the period 2000 to 2003.[1] Our research setting offers important advantages in examining the use of discretion in target setting. First, our data include initial store targets, which are calculated by applying an established formula to the regional targets set by corporate headquarters, and final store targets, which are set by regional directors. Any differences between the initial and the final targets allow us to draw inferences on how directors use discretion in the target-setting process. Second, in our setting, the sum of the store targets within each region must equal the total regional target set by corporate headquarters; that is, when a director adjusts one store's target downward, one or more other stores' targets must be adjusted upward. As a result, directors must be especially assiduous in

[1] We use the same research setting as Matsumura and Shin (2006), but we analyze a different dataset and a different time period.

using their discretion and, hence, the factors driving the use of discretion should be particularly informative in our setting. Moreover, the zero-sum nature of the adjustments creates a situation in which directors, after considering store managers' proposals, must choose which store managers to satisfy and which to displease. This allows us to examine the social construct of conflict avoidance that can arise in such a situation. Third, consistent with reported practice, store managers in our setting participate in the target-setting process by proposing targets and providing supporting documentation (see, e.g., Umapathy 1987; De With and Dijkman 2008; Shastri and Stout 2008). Our setting therefore allows us to expand the study of target setting beyond much of the prior literature that focuses on unilaterally determined targets (e.g., Locke and Latham 1990).

We find a positive and significant relation between downward target adjustments and our proxies for risk and fairness. These results support the view that directors strategically use their discretion in the target-setting process to manage compensation risk and mitigate fairness concerns. We further find a significantly positive relation between downward target adjustments and store managers' hierarchical status relative to both their director and other store managers. This evidence is consistent with directors using their discretion to avoid confrontations with higher-status store managers.

This study makes several important contributions to the literature. First, ours is among the first studies to take a broader perspective on the supervisor's role in the target-setting process. Prior research focuses on the relationship between target levels and employee effort, but does not take into account the fact that the target-setting process can also be strategically used to manage other elements essential in compensation contracting, such as compensation risk. In a setting in which standardized compensation contracts cover multiple units but there is some discretion in target levels, we find evidence suggesting that supervisors manage compensation risk by using their discretion in the target-setting process to compensate stores for additional store-specific risk.

Second, this study is the first to provide field data evidence consistent with supervisors exercising discretion in the target-setting process to counter other elements of the incentive contract that may be considered unfair. In particular, our results are consistent with supervisors using their discretion to improve fairness perceptions, even if the area in which the supervisor is granted discretion is not the source of the fairness concerns. Our results suggest that supervisors view different elements of the compensation contract as interrelated. When supervisors are not granted discretion to adjust every element of the compensation contract, discretion on one dimension is used to improve compensation contracting on other dimensions. More importantly, this evidence on the strategic use of discretion to harmonize elements of the compensation contract highlights the importance of studying the compensation system as a whole.

Third, this study contributes to the literature by examining the participative target-setting process from the supervisor's perspective. In contrast to much of the target-setting literature that focuses on motivating subordinates' decisions or effort levels, we shed light on how supervisors' individual incentives and preferences influence target-setting decisions. Results suggest that supervisors strategically make discretionary target adjustments to reduce confrontation costs. This implies that supervisors' personal preferences and incentives can result in not only biased performance evaluations at the end of the performance period, but also distorted target levels at the beginning of the contract period.

We describe our research setting in Section II and develop our hypotheses in Section III. In Section IV we discuss our sample and variables. Section V presents the results of our main analyses and Section VI presents the results of additional analyses and robustness tests. Section VII summarizes and concludes.

II. RESEARCH SETTING

Our research setting, which we refer to as POST, is the state-owned provider of comprehensive postal services in Korea. POST is a monopolistic service provider with respect to the handling

and delivery of regular mail. However, it competes against several large domestic and international couriers, such as Federal Express and DHL, in the expedited delivery market. Although expedited delivery services are subject to more intense competition, their volume is more controllable from a store manager's perspective because it can be influenced in part by store managers' actions. For instance, an individual store manager significantly increased the store's mail volume by targeting marketing efforts for express parcels to families in rural areas that were identified as having children who were a significant distance from home. Regular mail volume (mostly business bulk mail), on the other hand, is largely determined by factors beyond store managers' control. Because expedited delivery services constitute most of a store's "strategic mail services," we henceforth refer to expedited delivery services simply as strategic mail services.[2]

POST consists of eight regional headquarters, 214 Business Unit (BU) stores, and 2,579 small windows-only stores.[3] At the beginning of 2001, POST restructured the tasks and responsibilities of the store-level administrative staff to reduce the number of redundant administrative tasks performed. Some administrative tasks were delegated to other BU stores and some were moved to the marketing and sales department. At the same time, several BU stores merged in order to realize administrative efficiency gains. Although this did not reduce the number of store locations (the number of BU stores remained unchanged at 214), for administrative purposes POST now only recognizes 103 separate stores. We refer to these 103 higher-level BU stores as Evaluation Unit (EU) stores. Beginning in 2001, performance evaluation was performed at this more aggregated level (see Figure 1).

The Incentive Plan

POST's compensation system provides a fixed salary plus a bonus to all EU store employees, including the store manager. The fixed salary is based on individual qualifications and seniority, while the bonus depends on store performance. Bonuses range from 30 percent to 150 percent of the employee's monthly salary. This means that the maximum annual bonus constitutes approximately 12.5 percent of the employee's annual compensation. Management believes that this incentive system should motivate workers because such bonuses are unusually high among Korean state-owned companies.

As of 2001, the following three performance dimensions are measured and rewarded in POST's incentive plan: profitability, sales, and quality of service (see Table 1 for definitions). The incentive weights on the different dimensions are 60 percent, 20 percent, and 20 percent, respectively. The profitability dimension contains two equally weighted parts: profitability compared to the store's own profitability in the previous year and profitability compared to the profitability of a reference group. This relative performance evaluation (RPE) was introduced in 1998 in an effort to create a more fair reward system that takes into account various exogenous factors likely to affect performance.[4] Specifically, POST used cluster analysis to classify stores into reference groups of stores with similar business environments. In a continued effort to improve fairness, POST revised the reference groups in 1999. Nevertheless, in June 2000, POST recognized that some employees still perceived considerable unfairness in the reward system (Matsumura and Shin 2006).

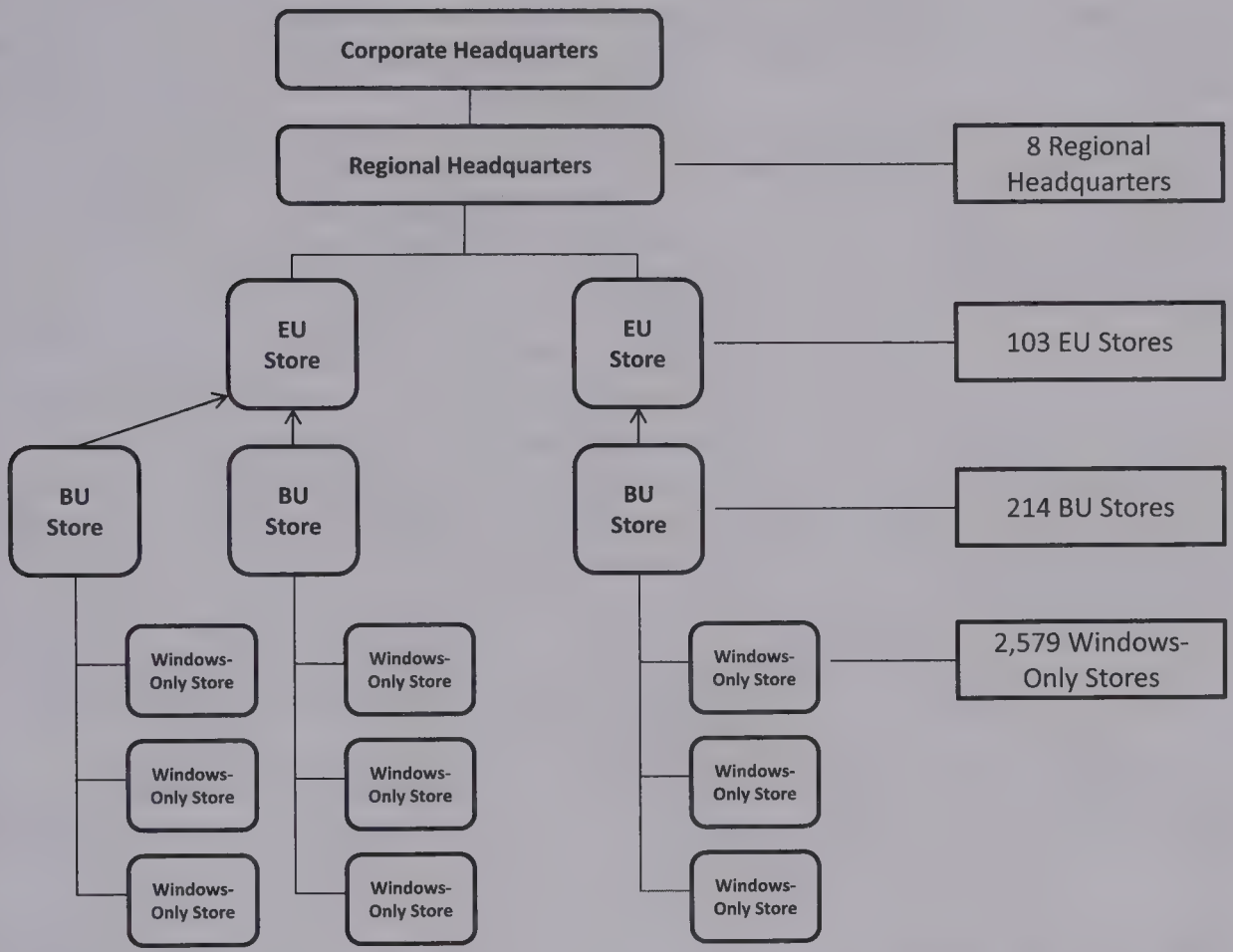
---

[2] POST's definition of strategic mail services includes domestic express mail, international express mail, domestic registered parcel, philately products, and congratulatory and condolence card services. According to an internal document prepared in 2003, POST estimates that their market shares of parcel and express mail services in the Korean market are 10.1 percent and 30.1 percent, respectively.

[3] We use the term Business Unit (BU) store to refer to stores that are equipped to perform the complete set of tasks (ranging from mail reception to final delivery), as opposed to small windows-only stores, which are only able to provide a limited set of services.

[4] Moreover, management also wanted to address gaming behavior that occurred in response to a previous incentive system that placed considerable weight on improvements from the previous year (Matsumura and Shin 2006).

FIGURE 1
Organizational Chart of POST



POST consists of 214 Business Unit (BU) stores and 2,579 small windows-only stores. For administrative and performance evaluation purposes, POST aggregates the BU stores and windows-only stores into 103 higher-level BU stores, which we refer to as Evaluation Unit (EU) stores.

The incentive plan's sales dimension contains two unequally weighted parts, one for regular mail services (with an incentive weight of 6 percent) and one for strategic mail services (with an incentive weight of 14 percent).[5] In both parts, the ratio of actual sales to target sales determines performance.

Finally, quality of service consists of three dimensions: management efficiency, customer satisfaction, and mail service quality. Customer satisfaction and mail service quality are measured

[5] One might wonder why sales are evaluated separately when a significantly large weight is already placed on profitability. The reason behind the separate measure is that the store's revenue used in calculating the profitability measure is "allocated" revenue, which is calculated by comparing the store's contribution to the entire value chain. The rationale is that mail service involves multiple steps, ranging from receiving and processing mail to final delivery, and these steps are usually performed by different stores that together form the value chain. In contrast, sales revenue simply refers to the revenue earned on sales that originate at the particular store.

TABLE 1

Description of Incentive Plan, 2001–2003

Performance Dimension	Performance Measure	Weight
Profitability	$\frac{\textit{Store Profitability}(t)}{\textit{Store Profitability}(t-1)}$	30%
	$\frac{\textit{Store Profitability}(t)}{\textit{Reference Group Profitability}(t)}$	30%
Sales Target	$\frac{\textit{Actual Sales of Regular Mail Service}(t)}{\textit{Target Sales of Regular Mail Service}(t)}$	6%
	$\frac{\textit{Actual Sales of Strategic Mail Service}(t)}{\textit{Target Sales of Strategic Mail Service}(t)}$	14%
Quality of Service	Management Efficiency	5%
	Customer Satisfaction	10%
	Mail Service Quality	5%

Store Profitability (t) = store-level allocated revenue divided by operating cost in year *t*;
Reference Group Profitability (t) = arithmetic mean of profitability of stores that belong to the specific reference group in year *t*;
Sales of Regular Mail Service (t) = sales revenue from regular mail volume received by the store at year *t*;
Sales of Strategic Mail Service (t) = sales revenue from strategic mail volume received by the store at year *t*.

by an independent outside agency contracted by corporate headquarters, and management efficiency is assessed by regional headquarters. All three dimensions receive high scores and the scores traditionally exhibit little variation across stores. POST's attention has therefore focused on increasing sales while maintaining cost effectiveness in order to increase profitability.

At the end of the year, a store's performance score is calculated using the performance dimensions and weights in the incentive plan (see Table 1). The stores are then ranked based on their performance scores. The store's rank determines the size of the bonus for each store. Hence, POST's incentive plan employs a two-stage RPE system that involves (1) RPE measures relative to a reference group of stores and (2) a tournament in which the size of the bonus depends on the store's performance score rank among all the stores. When the incentive plan was initially introduced in 1998, only the top 50 percent of the stores received a bonus. Because a large number of managers considered this unfair (Matsumura and Shin 2006), POST expanded the bonus eligibility to 100 percent of the stores in 2001. Although all stores now receive a bonus, the size of the bonus continues to depend on the store's rank.[6] Consequently, some of the weaker stores still feel that the system puts them at a disadvantage.

[6] POST also made other notable changes to the incentive plan introduced in 1998 (Matsumura and Shin 2006). The most relevant change from the perspective of this study is that as of 2001, POST replaced the cash flow measures with store-level sales targets. POST also eliminated the productivity dimension due to the high correlation between profitability and productivity measures. Instead, the weight placed on the profitability growth measure was significantly increased, leading to an equal weight of 30 percent on both the relative profitability and the profitability growth measures.

We focus this study on sales targets because these are the only major incentive plan design measures over which regional directors have discretion (described further below). While the 20 percent weight placed on the sales target measures might appear relatively low, we believe that more attainable sales targets are considered relevant by the stores because of POST's two-stage RPE system that ultimately bases the size of each store's bonus on its performance rank compared to other stores.[7]

The eight regional directors also receive a fixed salary and a bonus, where the bonus is determined by averaging the bonus rate (as a percent of monthly salary) of all stores under their direct supervision. Regional directors are not rewarded for meeting the regional sales target. By tying the regional directors' bonuses to their stores' bonus levels, POST ensures that the regional directors have incentives to consider how sales targets affect store managers' and workers' motivation.

The Target-Setting Process

The company's overall sales target plays an important role within POST, as the sales targets of the eight regional headquarters, and ultimately of the 103 EU stores, are based upon it. Responsibility for determining the company's overall sales target lies with a corporate task force team, which meets each November to set out the key strategic directions of the firm. Based on the overall macro-economic prospects, the business environment, and the intensity of competition, the team forecasts the coming year's mail volume, operating expenses, and sales. The proposed targets are reviewed by the CEO in December, and after the board of governors approves the targets, they serve as the formal financial targets for the coming year.

Using the company's overall financial targets, the corporate task force prepares operating and capital budgets for the eight regional headquarters. Targets for the regional headquarters are set in a top-down manner. In early January, the CEO presides over a target review committee meeting held at the corporate headquarters. In this meeting, the regional headquarters' budgets, including their sales targets, are presented to the regional directors and discussed in detail.

The regional directors then have responsibility for turning the total regional sales targets into store-specific sales targets. Although headquarters asks the regional directors to develop the initial targets based on a formula, the regional directors have discretion in the setting the targets. The only restriction headquarters imposes is that the sum of all individual stores' targets equals the regional sales target. Because the store managers have superior information about local market conditions, POST designed the target-setting process to be participative. Store managers can potentially influence the sales targets they receive by sharing their private information. Specifically, the steps in the target-setting process are as follows. First, the regional director privately determines the initial sales targets for each store by allocating the regional sales target to the stores according to the relative weight of the stores' prior sales. At the same time, without knowledge of the regional director's initial sales estimates, the store managers are requested to submit proposed sales targets for their stores, as well as supporting documentation to the regional director. When proposing sales targets and providing documentation, the store managers have no knowledge of the total regional sales targets or of the store-specific, formula-based initial targets.[8]

When the regional director has received all store managers' information, s/he compares the formula-based initial estimates to the targets proposed by the stores; based on this information the regional director arrives at his/her targets. In cases in which there is a large discrepancy between

[7] See, for example, Cichello et al. (2009) for evidence on the importance of small performance changes in tournament structures.

[8] The regional-level sales targets are shared with the store managers at the regional target review meeting in late January.

the director's target and the target the store proposes, the director solicits further information from the store before setting the final targets. By the end of January (about two weeks after the proposals are submitted), the regional director and all his/her store managers meet at the regional headquarters, and the regional director presents and discusses the sales targets.

III. HYPOTHESES

In the budgeting literature, considerable attention is directed to studying targets. The focus of many of these studies is on budgetary slack,[9] in particular, on how subordinates are able to obtain easier targets (i.e., create slack). The most robust finding from this literature is the relation between private information and slack. Subordinates are found to use their information advantage to negotiate performance targets that are more easily achieved (Schiff and Lewin 1968; Merchant 1985; Blanchard et al. 1986; Lukka 1988). In addition, when information asymmetry allows slack creation, subordinates' efforts to create slack are related to their marginal returns from effort. For example, Anderson et al. (2010) find that performance targets are lower when subordinates' career horizon and economic dependence on the job are greater. This relation suggests that subordinates expend greater effort to create slack when the returns from such effort are higher. Other factors that are relevant for slack creation include subordinates' risk preferences and preferences for honesty (e.g., Young 1985; Dunk and Nouri 1998; Evans et al. 2001).

Research that examines the target-setting process from the supervisor's side mainly focuses on how target achievability is related to subordinates' performance. The main prediction of this goal-setting literature is that there is a strong positive relation between target difficulty and performance (Locke and Latham 1990). Although the goal-setting literature consistently finds evidence indicating that difficult goals motivate better performance than easier goals, (e.g., Locke and Latham 1990), when Merchant and Manzoni (1989) re-examine these conclusions in a budgeting setting, they find that budget targets are more attainable than the goal-setting literature would predict. Interviews that Merchant and Manzoni (1989) conduct with managers suggest that targets are generally attainable because eliciting high levels of employee performance is not supervisors' only concern; their target-setting decisions are also affected by factors such as increasing the predictability of budgets and discouraging earnings management. Nevertheless, to date there exists little empirical evidence on the array of factors influencing supervisors' target-setting decisions and methods (Ittner and Larcker 2001; Anderson et al. 2010). As a result, an understanding of supervisors' target-setting behavior is far from complete.

We contribute to the knowledge on target setting by analyzing supervisors' target-setting decisions. Our study differs from Merchant and Manzoni (1989) in that we do not examine the general attainability of budget levels. Rather, in our setting, the levels of the regional sales targets, as well as the form of the incentive plan, are dictated by corporate headquarters, and we examine how supervisors use discretion to allocate the regional sales targets to specific stores. Specifically, we study how supervisors use discretion in the target-setting process to manage compensation risk, address fairness concerns, and minimize confrontation costs. Below we develop our predictions with respect to each of these three elements.

Compensation Risk

Much of principal-agent theory focuses on situations with outcome uncertainty, information asymmetry between the principal and agent, and compensation contracts based on measures that,

[9] Some notable exceptions are Murphy (2000) and Indjejikian and Nanda (2002), who study the performance standards that companies use to evaluate performance. Murphy (2000) examines the factors associated with the choice of an internal versus an external performance standard and Indjejikian and Nanda (2002) examine whether firms fully adjust performance standards for past performance.

due to uncertainty, are imperfectly informative about the agent's effort. Linking pay to performance induces the agent to expend effort, but also transfers some risk from the principal to the agent. The principal must compensate the agent for bearing this risk. The optimal compensation contract, therefore, trades off the costs and benefits of imposing compensation risk on the agent (Holmström 1979; Lambert 2001).

At POST, the form of stores' compensation contracts is standardized and set by corporate headquarters. When the regional directors determine each store's initial sales estimates based on the regional targets set by headquarters, the resulting contracts likely imposes excessive risk on specific stores' managers and employees. Because the regional directors cannot raise store managers' base salary to compensate for bearing more risk, we predict that directors will increase expected compensation to cover a risk premium by adjusting the sales targets downward for stores that face greater environmental uncertainty. This prediction is in line with prior budgeting research that finds a positive relation between slack (more attainable targets) and uncertainty (Merchant 1985; Govindarajan 1986; Schoute and Wiersma 2007).[10]

In sum, consistent with agency theory, we hypothesize that supervisors use their discretion in the target-setting process to manage store-specific compensation risk. More formally:

H1: Downward discretionary target adjustments are positively associated with store-specific compensation risk.

Fairness Concerns

Behavioral research shows that the perceived fairness of a compensation plan has a significant impact on employee incentives (e.g., Akerlof and Yellen 1988; Lindquist 1995; Konovsky 2000; Libby 2001). If an employee perceives his/her compensation outcome to be unfair, then s/he will have a negative behavioral reaction toward the compensation system and as a result will not be motivated by it (Cohen-Charash and Spector 2001; Colquitt et al. 2001). These negative behavioral reactions are especially strong when employees believe their outcomes would have been better if decision makers had implemented different procedures (Cropanzano and Folger 1989). Considering the significant influence on employee motivation, supervisors have clear incentives to take fairness concerns into account when managing a compensation plan.

Incentive compensation payouts at POST are partly based on a store's performance relative to its reference group (see Section II). Although the reference groups were created with the intention of increasing fairness in performance evaluation by grouping stores with similar operating environments, interview data reveal that some stores still consider the system unfair, as they believe that they are competing against much stronger stores (Matsumura and Shin 2006). Top management therefore remains concerned about dysfunctional behavior due to perceived unfairness and encourages regional directors to take fairness concerns seriously. Regional directors also have direct incentives to mitigate fairness concerns because decreased worker motivation impacts their own bonuses. Consequently, we expect regional directors to address fairness concerns. However, because they are unable to directly adjust the RPE element of the bonus system (the source of the unfairness complaints), we predict that they respond to these fairness issues by using the discretion they do have, which relates to the target-setting process. In particular, we predict that, by making sales targets easier to achieve for the stores facing more challenging reference groups,[11] supervi-

[10] Indjejikian and Matejka (2006) find no such relation.

[11] Recall that the sum of the targets must equal the regional target dictated by central headquarters, so reducing one store's target requires raising one or more other targets.

sors can make total incentive compensation more equitable in the sense of attainability. This prediction is in line with previous research that finds that more attainable targets help foster a sense of fairness (see, e.g., Merchant and Manzoni 1989).

In sum, we hypothesize that supervisors strategically use discretion in the target-setting process to help restore the perception of fairness among those managers and store employees who feel that the composition of the reference groups impairs the overall fairness of the incentive system. More formally:

H2: Downward discretionary target adjustments are positively associated with fairness concerns arising from other elements of the incentive system.

Conflict Avoidance

Supervisors are often nonresidual claimants whose incentives are not perfectly aligned with those of the company's owners. Hence, when supervisors tend to their own incentives in using their discretion during the compensation contracting process, their decisions will not necessarily be firm-value maximizing (e.g., Bol 2009). In a target-setting context, this suggests that some of the supervisors' target adjustments will reflect their own utility preferences, where utility encompasses both pecuniary and nonpecuniary elements. One way supervisors can reduce personal costs is by using their discretion in the target-setting process to avoid painful conflicts or confrontations, as these conflicts are stressful and hence costly to supervisors (Murphy and Cleveland 1995).

In our setting, where a downward adjustment to one initial target necessitates an upward adjustment in one or more other targets, conflict is inevitable. Nevertheless, supervisors can reduce confrontation costs by strategically determining which subordinates to disappoint and which to (partly) satisfy. By disappointing those subordinates who are less likely to confront their superior, supervisors avoid conflict as much as possible.

One factor influencing how much "push back" a supervisor will likely receive from a subordinate is the subordinate's status, that is, the amount of respect and prominence an individual holds within a social group (Anderson et al. 2001; Anderson and Berdahl 2002). Because of their social standing, high-status individuals are more likely to challenge their supervisor's decision. Consequently, communicating unfavorable information to relatively high-status individuals is expected to be more costly than communicating unfavorable information to relatively low-status individuals who are more likely to accept the decision without a challenge (Rahim 1986). Accordingly, we predict that supervisors are more likely to adjust targets downward for those subordinates who hold a relatively high status within the organization. This prediction is consistent with empirical research from psychology that suggests that managers resolve conflicts differently depending on the status of the individuals involved: supervisors tend to compromise with those who have similar status, whereas they tend to impose decisions on those with lower status (e.g., Rahim 1986; Phillips and Cheston 1979; Yukl and Falbe 1990). This prediction is also in line with findings that show that, consciously or not, individuals concede more to higher-status individuals in negotiations (Thye 2000; Thye et al. 2006).

We note that a related investigation in the accounting domain is Anderson et al. (2010), who examine whether job tenure influences target attainability. The authors find a negative association between subordinates' job tenure and target levels.[12] Although job tenure could proxy for status, Anderson et al. (2010) argue that in their setting, the negative association between job tenure and target levels is more likely the result of bargaining power from increased local knowledge, not status. In contrast, the source of status that is especially pertinent in our setting is job grade within

[12] Indjejikian and Matejka (2006) incorporate business-unit controller job tenure as a control variable but find no evidence that it is related to more attainable targets.

the organization's hierarchy. This source is not correlated with increased knowledge of the local environment, as POST maintains a policy of frequent job rotation.[13] Hence, supervisors' consideration of hierarchical status arises from individual influence, not from private information.

In sum, we predict that supervisors use their discretion in the target-setting process to decrease the sales targets of high-status store managers to reduce confrontation costs. More formally:

H3: Downward discretionary target adjustments are positively associated with higher relative hierarchical status.

IV. RESEARCH DESIGN

Sample

To study the target-setting process, we collected annual performance evaluation data from 2000 (the first year sales targets were part of the incentive plan) through 2003 for POST's 103 EU stores. We obtained financial and operational data of individual stores from POST's performance measurement system, and control variables such as demographics and individual store characteristics from one of the company's databases that tracks information on individual stores. To gain an understanding of the incentive plan and the company's background, we studied a variety of company documents and interviewed corporate headquarters' managers, including one of POST's incentive plan designers.

Dependent Variable

To capture the regional directors' use of discretion in the target-setting process, we measure the extent to which regional directors adjust the initial formula-based sales estimate. Recall from Section II that the regional directors allocate the regional sales targets set by corporate headquarters into store-level sales targets. The regional director first calculates the initial sales estimate by allocating the regional sales target according to the relative weight of the stores' prior-year sales and then, after receiving information from the individual stores, adjusts this estimate upward or downward to arrive at the final sales target. We use the size of the adjustment (*ADJUST*), measured by subtracting the final store-level sales target (*TAR_SALES*) from the regional director's initial formulaic estimate (*TAR_SALES_SUP*), as our proxy for regional directors' use of discretion at each store. We designate *ADJUST_REG* and *ADJUST_STRA* as the adjustments for regular sales and strategic sales, respectively.

Explanatory Variables

Compensation Risk Due to Environmental Uncertainty

To examine whether regional directors use discretion in the target-setting process to manage compensation risk, we identify stores that face different levels of compensation risk. We do so by capturing store-specific levels of environmental uncertainty. In our setting, regular and strategic mail volumes received by a store are summary measures that capture external changes in the local business environment. These measures allow us to use the change in regular and strategic mail

[13] We contacted the 103 stores directly and requested store managers' tenure data (as of January of 2001, 2002, and 2003, when the target-setting process occurs). We received 218 store-year managers' tenure data (out of 309 possible observations). The mean (median) store manager tenure was 11 (9) months and the maximum store manager tenure was 35 months. We also obtained a list of the 22 former regional directors who served at a specific regional headquarters from 1970 to 2003 and found that the mean tenure of a regional director was 19 months. The short tenure of store managers and regional directors results from the traditional rank-in-person system of the Korean government, where one is hired not for a specific position but as an individual suited to occupy multiple positions. This system has been criticized for limiting employees from developing expertise in specialized fields. However, this is not the case for POST's store managers because only the location and generally not the job function changes.

volume to capture the volatility of the business environment faced by each store. More specifically, we use the time-series standard deviation of the logarithm of regular (strategic) mail volume received by a particular store over the most recent four years to create *NOISE_MV_REG* (*NOISE_MV_STRA*) (Holthausen et al. 1995; Nagar 2002).

Fairness Concerns

To analyze whether store managers' concerns about the fairness of the incentive plan influence regional directors' use of discretion in the target-setting process, we create a proxy for store managers' fairness perceptions. As described in Section II, one of POST's main performance measures is based on a comparison to a reference group of similar stores. Matsumura and Shin (2006) document that store managers who feel that their reference group is too challenging are not motivated to exert effort because they consider the incentive system unfair. We create a proxy that captures such perceptions of unfairness by measuring how challenging a store's reference group is. To capture the difficulty of the reference group, we examine a store's profitability as compared to its reference group in the prior year. However, because weak performance relative to the reference group could be due to reasons other than a challenging reference group (e.g., poor management or lack of effort), we also incorporate the store's performance compared to the prior year's fixed sales targets. We reason that weak performance relative to the reference group but strong performance compared to the individual target indicates a challenging reference group.

Consistent with Matsumura and Shin (2006), we define *UNFAIR_REG* as the difference between a store's yearly rank on its RPE profitability measure (profitability divided by the reference group mean) and the store's yearly rank on the regular sales target measure in the prior year. That is:

$$UNFAIR\_REG = Rank\left[\frac{Store\ Profitability_{t-1}}{Reference\ Group\ Profitability_{t-1}}\right] - Rank\left[\frac{Actual\ Sales\ of\ Regular\ Mail_{t-1}}{Target\ Sales\ of\ Regular\ Mail_{t-1}}\right].$$

UNFAIR_STRA is defined similarly for strategic mail, using the actual and target sales of strategic mail. Higher values of the *UNFAIR* variables indicate higher unfairness.[14] That is, we expect an increased perception of unfairness when a store's profitability relative to its peers is low even though it meets or beats its sales targets.

Conflict Avoidance

To examine whether directors use discretion in target setting to reduce confrontation costs by avoiding conflicts with relatively higher-status store managers, we use job grade level, which is one of the most important indicators of status at POST.[15] All BU-level store managers of POST are level 4 or 5 employees and regional directors are level 1, 2, or 3 employees of the Korean government's pay schedule. We use this information to create the indicator variable *STATUS*,

[14] Thus, the possible rank differences range from −102 to 102. Suppose store A is ranked 93rd (out of 103 stores) on the RPE profitability measure but ranked 5th on the regular sales target measure in 2000. This indicates that while store A's actual sales far exceeded its regular mail sales target, its profitability relative to peer stores is very low. The difference in ranking for this store is 88, resulting in a value of 88 on our *UNFAIR_REG* variable. In contrast, suppose store B is ranked 1st on the RPE profitability measure but ranked 98th on the regular sales target measure in 2002. Store B's difference in ranking is −97, resulting in a value of −97 on our *UNFAIR_REG* variable.

[15] Similar to the General Schedule of U.S. Federal employees, the Korean government uses job grades that conform to the government's pay schedule. The pay schedule assigns a grade level from 10 to 1 (with 1 the highest) to each job according to the minimum level of education and experience that the position requires. An employee's salary increases as that employee's level on the scale increases. Typically, an employee moves up the grade scale as his/her seniority

which equals 1 if the difference in job grade level between the regional director and the store manager is 2 or less, and 0 otherwise.[16] Hence, *STATUS* is 1 if the status difference between the regional director and the store manager is small.

Control Variables

We include a set of proxy variables that control for differences in target-setting behavior. First, we control for the level of information asymmetry between regional directors and store managers, as the budgeting literature has long recognized the potential effect of information asymmetry in a participative budgeting process (e.g., Schiff and Lewin 1970; Christensen 1982; Baiman and Evans 1983; Indjejikian and Matejka 2006). Drawing on the performance evaluation literature, we identify divisional complexity as an important source of information asymmetry (Holthausen et al. 1995; Bushman et al. 1995; Keating 1997; Bushman et al. 2004; Berry et al. 2006). We capture the complexity of the different EU stores by including the number of small windows-only stores (*N_SUB*) and the number of BU stores added to an EU store for performance evaluation purposes since 2001 (*N_MERGE*).[17]

Second, we include the number of store employees (*N_EMP*) as a proxy for size (Indjejikian and Matejka 2006). Because the status variable is likely correlated with store size, it is important to control for size to eliminate potential confounding effects.

Third, we control for the competitive environment of the area (*COMPET*) in which the store operates, as prior research documents that the degree of competition influences budget misrepresentation (Bruggen and Luft 2009). The variable *COMPET* takes the value of 1 for urban areas and 0 for rural areas. We distinguish between urban and rural areas because major competitors of POST, such as domestic and international couriers, mainly operate in profitable urban areas (see Banker and Mashruwala [2007] for a similar distinction in a retail industry setting). This variable also captures the economic viability of the area in which the store operates.

Finally, in addition to year indicator variables, we include seven region indicator variables (*REGION*) to control for unspecified supervisor and region-specific differences in the target-setting process.

The above discussion leads to the following specifications:

$$\begin{aligned} ADJUST\_REG_{i,t} = {} & \alpha_0 + \alpha_1 NOISE\_MV_{i,t-1} + \alpha_2 UNFAIR_{i,t-1} + \alpha_3 STATUS_{i,t-1} + \alpha_4 N\_SUB_{i,t-1} \\ & + \alpha_5 N\_MERGE_{i,t-1} + \alpha_6 N\_EMP_{i,t-1} + \alpha_7 COMPET_{i,t-1} \\ & + REGION\_INDICATORS + YEAR\_INDICATORS + \varepsilon_{i,t} \end{aligned} \quad (1)$$

$$\begin{aligned} ADJUST\_STRA_{i,t} = {} & \alpha_0 + \alpha_1 NOISE\_MV_{i,t-1} + \alpha_2 UNFAIR_{i,t-1} + \alpha_3 STATUS_{i,t-1} + \alpha_4 N\_SUB_{i,t-1} \\ & + \alpha_5 N\_MERGE_{i,t-1} + \alpha_6 N\_EMP_{i,t-1} + \alpha_7 COMPET_{i,t-1} \\ & + REGION\_INDICATORS + YEAR\_INDICATORS + \varepsilon_{i,t} \end{aligned} \quad (2)$$

where Equation (1) pertains to regular mail, Equation (2) pertains to strategic mail, i indicates

within the firm increases, although above-average job performance can accelerate this progress. We obtained store managers' salary data for 1997 and found that the average salary for a level 4 store manager is significantly higher than that of a level 5 store manager ($p = 0.07$).

[16] The mean (median) job grade difference between a store manager and a regional director in our sample is 2.18 (2.00). Thus, we use the median job grade difference (2) as our cut-off to create the indicator variable *STATUS*.

[17] Prior research also documents budget emphasis and the degree of budget participation as factors that influence subordinates' target-setting behavior (e.g., Merchant 1985; Dunk 1993; Dunk and Nouri 1998). However, because the emphasis on targets and the degree of participation in the target-setting process are consistent over our sample, we do not separately control for these factors.

stores (i = 1,...,103), and t indicates time (t = 2001, 2002, 2003).[18] We estimate Equations (1) and (2) using OLS regression analysis with Huber-White robust standard errors clustered by store. These standard errors are robust to both serial correlation and heteroscedasticity (Rogers 1993).[19]

V. RESULTS

Descriptive Statistics and Univariate Correlations

Table 2 presents descriptive statistics of the variables used in the analyses.[20] The average store-level sales target for regular mail services (*TAR_SALES_REG*) is 7.81 billion Korean Won, while the average sales target for strategic services (*TAR_SALES_STRA*) is 2.27 billion Korean Won. The mean of *ADJUST* is 0 by design because the target-setting process at the regional level is a zero-sum game. The mean of *NOISE_MV_STRA* is higher than that of *NOISE_MV_REG*, which indicates that strategic mail volume displays higher temporal variability than regular mail volume. The mean of *UNFAIR* is also 0 by construction because this measure is based on the raw difference between a store's rank on the relative profitability measure and its rank on the sales target measure. The mean of *STATUS* is 0.57; that is, 57 percent of the store managers have a job grade that is within 2 levels of their regional supervisor. On average, EU stores are responsible for approximately 25 windows-only stores and two BU stores (the mean of *N_MERGE* is 1.07). While not reported in Table 2, regional headquarters are responsible for supervising 22 EU-level stores on average. Consistent with Merchant and Manzoni's (1989) field-based findings that targets are set to be achievable 80 to 90 percent of the time, POST stores' actual sales in regular (strategic) mail exceed target sales 73 percent (83 percent) of the time.[21] Table 3 provides details on Pearson correlations among the variables used in the analyses.

Determinants of Supervisors' Target Adjustments

Columns (1) and (2) of Table 4 present the estimation results of Equations (1) (regular mail) and (2) (strategic mail), which examine regional directors' use of discretion in the target-setting process. Consistent with H1, the coefficient on *NOISE_MV*, our proxy for environmental uncertainty, is positive and significant at the 5 percent level for strategic mail. This is consistent with regional directors using their discretion in the target-setting process to manage compensation risk arising from environmental uncertainty. The coefficient on *NOISE_MV* is not significant for regular mail. A possible explanation for this finding is that the lower differential compensation risk for regular mail (the temporal variability of mail volume is lower for regular mail than for strategic mail, as seen in Table 2) reduces the need for risk adjustments to a level where our tests are not powerful enough to detect them empirically. The low environmental uncertainty in conjunction with the much lower weight placed on the regular mail sales target (i.e., 6 percent for regular mail

---

[18] Because the BU store-level sales targets are determined at the beginning of the year, all explanatory variables are lagged values.

[19] Because the sum of all store-level adjustments that a regional director makes in a given year must be zero, we check whether our findings are robust to potential error dependence among stores in the same region for a given year by using standard errors clustered by 24 region-year pairings (8 regions for 3 years). In addition, we run yearly regressions of Equations (1) and (2) using standard errors clustered by region. The results are similar to those reported in Table 4.

[20] Descriptive statistics are based on 309 store-year observations from 2001 to 2003 for multivariate analyses.

[21] We compare the relative forecast accuracy of the final targets (*TAR_SALES*) and the director's initial formula-based estimate (*TAR_SALES_SUP*) with respect to predicting actual sales. Consistent with the measurement of analyst forecast error in financial archival research, we construct a measure of forecast accuracy by taking the absolute value of the difference between the final sales target (initial sales estimate) and actual sales, scaled by actual sales. We find that the regional director's initial formula-based estimate is a more accurate predictor (i.e., exhibits smaller forecast error) of actual sales for both regular and strategic mail. The paired t-tests suggest that the difference in forecast accuracy is statistically significant at the 1 percent level. This finding lends support to the conjecture that supervisors' strategic and self-serving use of discretion in target setting may result in targets that are less predictive of future performance.

TABLE 2

Descriptive Statistics

Variable	Mean	Std. Dev.	Q1	Median	Q3	n
$TAR\_SALES\_REG_t$	7.81	7.29	3.84	5.97	8.67	309
$TAR\_SALES\_STRA_t$	2.27	1.86	1.29	1.83	2.48	309
$ACT\_SALES\_REG_t$	8.42	7.38	4.01	6.46	9.63	309
$ACT\_SALES\_STRA_t$	2.46	1.78	1.49	2.02	2.84	309
$TAR\_SALES\_REG\_SUP_t$	7.81	6.83	3.72	6.01	9.05	309
$TAR\_SALES\_STRA\_SUP_t$	2.27	1.70	1.31	1.83	2.61	309
$ADJUST\_REG_t$	0.00	1.15	−0.23	−0.03	0.28	309
$ADJUST\_STRA_t$	0.00	0.32	−0.08	0.02	0.12	309
$NOISE\_MV\_REG_{t-1}$	0.11	0.13	0.04	0.07	0.12	309
$NOISE\_MV\_STRA_{t-1}$	0.16	0.11	0.09	0.14	0.20	309
$UNFAIR\_REG_{t-1}$	0.00	40.79	−26.00	−2.00	29.00	309
$UNFAIR\_STRA_{t-1}$	0.00	40.84	−28.00	1.00	26.00	309
$STATUS_{t-1}$	0.57	0.50	0.00	1.00	1.00	309
N_SUB_{t-1}	25.11	13.93	14.00	22.00	35.00	309
N_MERGE_{t-1}	1.07	1.01	0.00	1.00	2.00	309
N_EMP_{t-1}	2.50	0.86	1.88	2.43	3.05	309
$COMPET_{t-1}$	0.61	0.49	0.00	1.00	1.00	309

Variable Definitions:

TAR_SALES_REG = store-level sales target for regular mail (in billion Korean Won);
TAR_SALES_STRA = store-level sales target for strategic mail (in billion Korean Won);
ACT_SALES_REG = store-level actual sales from regular mail (in billion Korean Won);
ACT_SALES_STRA = store-level actual sales from strategic mail (in billion Korean Won);
TAR_SALES_REG_SUP = regional director's initial sales estimate for regular mail (in billion Korean Won);
TAR_SALES_STRA_SUP = regional director's initial sales estimate for strategic mail (in billion Korean Won);
ADJUST_REG = *TAR_SALES_REG_SUP – TAR_SALES_REG;*
ADJUST_STRA = *TAR_SALES_STRA_SUP – TAR_SALES_STRA;*
NOISE_MV_REG = time-series standard deviation of the logarithm of regular mail volume received by the store over the most recent four years;
NOISE_MV_STRA = time-series standard deviation of the logarithm of strategic mail volume received by the store over the most recent four years;
UNFAIR_REG = difference between the store's rank on relative profitability and rank on the sales target measure for regular mail, with higher values indicating stores that exhibited a high achievement of regular mail sales compared to their target, but low relative profitability;
UNFAIR_STRA = difference between the store's rank on relative profitability and rank on the sales target measure for strategic mail, with higher values indicating stores that exhibited a high achievement of strategic mail sales compared to their target, but low relative profitability;
STATUS = indicator variable that takes the value of 1 if the difference in job grade level between the regional director and the store manager is 2 or less, and 0 otherwise;
N_SUB = number of small windows-only stores the EU manager is responsible for;
N_MERGE = number of BU stores added to an EU store for performance evaluation purposes;
N_EMP = number of employees (in hundreds); and
COMPET = indicator variable that takes the value of 1 for stores in urban areas and 0 for stores in rural areas.

versus 14 percent for strategic mail) make it less likely that the costs of making the adjustment outweigh the benefits.

TABLE 3

Pearson Correlations among Variables

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. $ADJUST\_REG_t$	1.00										
2. $ADJUST\_STRA_t$	0.33	1.00									
3. $NOISE\_MV\_REG_{t-1}$	0.09^{NS}	0.07^{NS}	1.00								
4. $NOISE\_MV\_STRA_{t-1}$	0.09^{NS}	0.10*	0.49	1.00							
5. *STATUS*	0.02^{NS}	-0.01^{NS}	−0.11*	−0.15	1.00						
6. $UNFAIR\_REG_{t-1}$	0.28	0.15	0.06^{NS}	0.10*	−0.11**	1.00					
7. $UNFAIR\_STRA_{t-1}$	0.25	0.23	-0.11^{NS}	0.00^{NS}	0.06^{NS}	0.41	1.00				
8. N_SUB_{t-1}	0.08^{NS}	0.13**	−0.21	−0.22	0.44	-0.02^{NS}	0.11 *	1.00			
9. N_MERGE_{t-1}	0.08^{NS}	0.15	−0.22	−0.25	0.38	0.00^{NS}	0.10 *	0.84	1.00		
10. N_EMP_{t-1}	−0.13**	−0.13**	−0.24	−0.17	0.08^{NS}	-0.08^{NS}	-0.06^{NS}	0.76	0.67	1.00	
11. $COMPET_{t-1}$	-0.01^{NS}	-0.08^{NS}	0.21	0.30	−0.33	0.02^{NS}	-0.08^{NS}	−0.40	−0.31	-0.05^{NS}	1.00

*,**Significantly different from 0 at the 0.10 and 0.05 percent level, respectively (two-tailed).

^{NS} Not significantly different from 0 at the 0.10 level (two-tailed).

All other correlation coefficients are significantly different from 0 at the 0.01 percent level (two-tailed).

For variable definitions, see Table 2.

TABLE 4

Determinants of Supervisors' Target Adjustments (t = 2001–2003)

$$ADJUST_{i,t} = \alpha_0 + \alpha_1 NOISE\_MV_{i,t-1} + \alpha_2 UNFAIR_{i,t-1} + \alpha_3 STATUS_{i,t-1} + \alpha_4 N\_SUB_{i,t-1} + \alpha_5 N\_MERGE_{i,t-1} + \alpha_6 N\_EMP_{i,t-1} + \alpha_7 COMPET_{i,t-1} + REGION\_INDICATORS + YEAR\_INDICATORS + \varepsilon_{i,t}$$

		(1)	(2)
		ADJUST	
		ADJUST_REG	*ADJUST_STRA*
Variable	**Predicted Sign**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
Constant	?	0.944**	0.378***
		(2.15)	(2.95)
$NOISE\_MV_{t-1}$	(+)	0.154	0.305**
		(0.30)	(2.02)
$UNFAIR_{t-1}$	(+)	0.006***	0.001*
		(2.64)	(1.40)
$STATUS_{t-1}$	(+)	0.604***	0.107**
		(2.95)	(1.65)
N_SUB_{t-1}	(+)	0.069***	0.022***
		(3.47)	(3.30)
N_MERGE_{t-1}	(+)	0.143	0.097**
		(1.09)	(2.92)
N_EMP_{t-1}	?	−1.024***	−0.357***
		(−4.65)	(−3.45)
$COMPET_{t-1}$	?	0.351**	0.053
		(2.11)	(1.09)
n		309	309
Adjusted R^2		23.5%	31.6%

*, **, *** Significant at the 0.10, 0.05, and 0.01 levels, respectively (one-tailed test if a signed prediction is provided, two-tailed otherwise).

For parsimony, coefficients on region and year indicators are not reported.

Reported t-statistics are based on the Rogers (1993) standard errors.

Variable Definitions:

ADJUST = *ADJUST_REG* or *ADJUST_STRA;*
ADJUST_REG = *TAR_SALES_REG_SUP – TAR_SALES_REG*;
ADJUST_STRA = *TAR_SALES_STRA_SUP – TAR_SALES_STRA*;
NOISE_MV_REG = time-series standard deviation of the logarithm of regular mail volume received by the store over the most recent four years;
NOISE_MV_STRA = time-series standard deviation of the logarithm of strategic mail volume received by the store over the most recent four years;

(continued on next page)

TABLE 4 (continued)

UNFAIR_REG = difference between the store's rank on relative profitability and rank on the sales target measure of regular mail, with higher values indicating stores that exhibited a high achievement of regular mail sales compared to their target, but low relative profitability;
UNFAIR_STRA = difference between the store's rank on relative profitability and rank on the sales target measure of strategic mail, with higher values indicating stores that exhibited a high achievement of strategic mail sales compared to their target, but low relative profitability;
STATUS = indicator variable that takes the value of 1 if the difference in job grade level between a regional director and a store manager is 2 or less, and 0 otherwise;
N_SUB = number of small windows-only stores the EU manager is responsible for;
N_MERGE = number of BU stores added to an EU store for performance evaluation purposes since 2001;
N_EMP = number of employees (in hundreds); and
COMPET = indicator variable that takes the value of 1 for stores in urban areas and 0 for stores in rural areas.

Next, as predicted by H2, the results in Table 4 suggest that regional directors use their discretion to address apparent unfairness of other elements of the incentive contract. The coefficients on *UNFAIR* are positive and significant at the 1 and 10 percent levels for regular and strategic mail, respectively. Recall from Section II that from a store manager's perspective, regular mail is less controllable than strategic mail. This difference is relevant, as competing against a reference group of stronger stores will likely be even more frustrating when the manager has little control over the specific dimension. When addressing fairness concerns, regional directors are therefore more likely to adjust targets on the less controllable dimension than on the more controllable dimension. Hence, the stronger result for regular mail, which is less controllable, is consistent with our fairness-based predictions.[22]

Table 4 also shows that, consistent with H3, the coefficients on *STATUS* are positive and significant at the 1 percent level for regular mail and at the 5 percent level for strategic mail. This finding suggests that supervisors assign more attainable targets to store managers of more similar social status to reduce confrontation costs.

Finally, our results corroborate the findings of prior literature on budgetary slack, suggesting that complexity of operations creates information asymmetry that subordinates use to their advantage to obtain more attainable targets. Specifically, the coefficients on *N_SUB* are positive and significant at the 1 percent level for both regular and strategic mail and the coefficient on *N_MERGE* is positive and significant at the 5 percent level for strategic mail.

[22] As discussed in Section II, a regional director's bonus rate is the average bonus rate of all stores in the same region. Hence, assigning more (less) attainable sales targets to stores when poor (favorable) RPE performance is anticipated may also be consistent with a regional director's economic incentive. If a regional director's economic incentive drives our results, however, we would expect stronger results for strategic mail than for regular mail because strategic mail receives a higher incentive weight (14 percent as compared to 6 percent of regular mail) and therefore provides an easier way to create an impact. This is not what the results show (see Table 4). Furthermore, if a regional director solely tries to maximize his/her bonus payout when setting store-level sales targets, s/he would only consider a store's prior year performance on the RPE incentive element rather than comparing the store's RPE performance to its sales target performance. When we replace *UNFAIR_REG* and *UNFAIR_STRA* with a store's rank on relative profitability in the prior year to better capture a director's self-interest, we find that the coefficient on RPE rank is insignificant ($p = 0.13$ (0.35) one-tailed) for regular (strategic) mail in contrast to the results reported in Table 4. Overall, we conclude that our results are more consistent with our fairness-based predictions.

VI. ADDITIONAL ANALYSES AND ROBUSTNESS TESTS

Related Dependent Variables: Attainability

In the main analysis, we use the extent to which the regional directors adjust their formulaic initial estimate upward or downward for final target determination as a proxy for supervisors' use of discretion in target setting. Because of the focus in the budgeting literature on the related issue of target attainability, and because target attainability is one of the factors that regional directors consider as they adjust initial targets upward or downward, we compare our findings to results using two measures of target attainability as the dependent variable.

First, in the target-setting process, supervisors likely make a private estimate of future store sales that, unlike the initial sales estimate, is based not only on past sales, but also on the regional director's knowledge of the store's characteristics and competitive environment. We try to capture this private estimate by using a prediction model that includes past sales, store characteristics, and environmental characteristics available in POST's databases (see the Appendix for details). We use the store-year-level predicted values of sales resulting from the prediction model to proxy for the regional directors' private estimates of future sales. We then compute target attainability (*ATTAIN1*) by taking the difference between the proxy for the directors' estimates of sales and the final sales targets. A positive (negative) value means that the target is set lower (higher) than the level predicted by the exogenous factors influencing the store-level sales, which we interpret as implying that the sales target is more (less) attainable.

Second, following prior studies (e.g., Merchant and Manzoni 1989; Indjejikian and Nanda 2002; Indjejikian and Matejka 2006; Anderson et al. 2010) we use the difference between actual sales and target sales as an additional proxy for target attainability (*ATTAIN2*).

Table 5 presents the results from estimating Equations (1) and (2) when we use our proxies for target attainability, *ATTAIN1* and *ATTAIN2*.[23] In contrast to the results of our main analysis concerning H1, the coefficient on *NOISE_MV* is positive and significant at the 1 percent level for regular mail but is not significant for strategic mail when *ATTAIN1* is used as a dependent variable. When we use *ATTAIN2* as a dependent variable, however, the results are similar to those reported in Table 4.

Consistent with the results of our main analysis presented in Table 4, we find support for H2 and H3 when using the alternative dependent variables. The coefficient on *UNFAIR* is positive and significant at the 5 percent level or less except in column (4), where the coefficient is not significant ($p = 0.46$, one-tailed). The coefficient on *STATUS* is also positive and significant at the 5 percent level except in column (4), where the coefficient is not significant ($p = 0.12$, one-tailed). We also confirm the effect of information asymmetry on target attainability. In all four regressions, the coefficients on *N_SUB* are positive and significant. Overall, we are able to confirm the findings from our main analysis when we replace the extent of regional directors' discretionary target adjustments with measures of target attainability,[24] suggesting that target attainability appears to be a key factor underlying regional directors' target adjustments. This result is also consistent with findings in prior research that supervisors consider a number of factors when they determine business units' targets (Merchant and Manzoni 1989).

---

[23] As expected, our proxies for target attainability are positively correlated with our measure of supervisors' target adjustments (*ADJUST*) at the 1 percent level. The correlation between *ADJUST_REG* (*STRA*) and *ATTAIN1_REG* (*STRA*) is 0.30 (0.36) and the correlation between *ADJUST_REG* (*STRA*) and *ATTAIN2_REG* (*STRA*) is 0.64 (0.81).

[24] By comparing actual performance against target performance over time, prior research documents that exceeding the target in the current year is positively associated with the probability of exceeding the target in the next year (Indjejikian and Nanda 2002; Indjejikian and Matejka 2006). This suggests that our measures of target attainability, and *ATTAIN2* in particular, could be serially correlated. To control for potential serial correlation we include the lagged value of *ATTAIN1* and *ATTAIN2* in the estimation models; this does not change our inferences.

TABLE 5

Determinants of Target Attainability
(t = 2001–2003)

$$ATTAIN_{i,t} = \alpha_0 + \alpha_1 NOISE\_MV_{i,t-1} + \alpha_2 UNFAIR_{i,t-1} + \alpha_3 STATUS_{i,t-1} + \alpha_4 N\_SUB_{i,t-1} + \alpha_5 N\_MERGE_{i,t-1} + \alpha_6 N\_EMP_{i,t-1} + \alpha_7 COMPET_{i,t-1} + REGION\_INDICATORS + YEAR\_INDICATORS + \varepsilon_{i,t}$$

		(1)	(2)	(3)	(4)
		Alternative Dependent Measure			
		ATTAIN1		*ATTAIN2*	
	Predicted Sign	*REG*	*STRA*	*REG*	*STRA*
Constant	?	0.178	0.738**	0.884**	0.235**
		(0.49)	(2.16)	(2.20)	(2.36)
$NOISE\_MV_{t-1}$	(+)	1.646***	0.531	0.429	0.352**
		(2.48)	(1.26)	(0.62)	(2.17)
$UNFAIR_{t-1}$	(+)	0.012***	0.005***	0.006**	0.000
		(4.99)	(3.38)	(1.97)	(0.10)
$STATUS_{t-1}$	(+)	0.708**	0.399**	0.292**	0.064
		(2.13)	(2.14)	(1.99)	(1.18)
N_SUB_{t-1}	(+)	0.067***	0.038***	0.034**	0.017***
		(3.07)	(2.38)	(2.09)	(3.41)
N_MERGE_{t-1}	(+)	0.013	0.051	−0.171	0.060**
		(0.09)	(0.61)	(−1.43)	(2.12)
N_EMP_{t-1}	?	−0.655**	−0.701***	−0.129	−0.210***
		(−2.13)	(−2.60)	(−1.05)	(−2.85)
$COMPET_{t-1}$	?	1.161***	0.480***	0.450***	0.068*
		(6.28)	(3.86)	(2.97)	(1.51)
n		309	309	309	309
Adjusted R^2		42.3%	23.4%	32.3%	24.3%

*, **, *** Significant at the 0.1, 0.05, and 0.01 levels, respectively (one-tailed test if a signed prediction is provided, two-tailed otherwise).

For parsimony, coefficients on region and year indicators are not reported.

Reported t-statistics are based on the Rogers (1993) standard errors.

Variable Definitions:

ATTAIN = *ATTAIN_REG or ATTAIN_STRA;*
ATTAIN1_REG = store and year-specific predicted value obtained by estimating Equation (3) in the Appendix minus *TAR_SALES_REG*;
ATTAIN1_STRA = store and year-specific predicted value obtained by estimating Equation (4) in the Appendix minus *TAR_SALES_STRA*;
ATTAIN2_REG = *ACT_SALES_REG – TAR_SALES_REG*;
ATTAIN2_STRA = *ACT_SALES_STRA – TAR_SALES_STRA*;
NOISE_MV_REG = time-series standard deviation of the logarithm of regular mail volume received by the store over the most recent four years;

(continued on next page)

TABLE 5 (continued)

NOISE_MV_STRA = time-series standard deviation of the logarithm of strategic mail volume received by the store over the most recent four years;
UNFAIR_REG = difference between the store's rank on relative profitability and rank on the sales target measure for regular mail, with higher values indicating stores that exhibited a high achievement of regular mail sales compared to their target, but low relative profitability;
UNFAIR_STRA = difference between a store's rank on relative profitability and rank on the sales target measure for strategic mail, with higher values indicating stores that exhibited a high achievement of strategic mail sales compared to their target, but low relative profitability;
STATUS = indicator variable that takes the value of 1 if the difference in job grade level between the regional director and the store manager is 2 or less, and 0 otherwise;
N_SUB = number of small windows-only stores the EU manager is responsible for;
N_MERGE = number of BU stores added to an EU store for performance evaluation purposes since 2001;
N_EMP = number of employees (in hundreds); and
COMPET = indicator variable that takes the value of 1 for stores in urban areas and 0 for stores in rural areas.

Robustness Tests

Alternative Measures of Environmental Uncertainty

To test the robustness of our findings related to compensation risk due to environmental uncertainty, we capture environmental uncertainty in alternative ways. Specifically, in line with Nagar (2002), we use the average percentage growth rate in regular and strategic mail volume received by a store over the past three years and the prior year's percentage growth rate. All inferences remain unchanged when using these alternative measures.

Alternative Measures of Unfairness

We explore the robustness of our findings related to fairness concerns by re-estimating Equations (1) and (2) using alternative measures of unfairness. First, we use an indicator variable that takes the value of 1 if the rank difference between the RPE profitability measure and the sales target measure is in the top 10 percent (50 percent), and 0 otherwise. Second, we take the difference between the store's yearly rank on its RPE profitability measure and the store's yearly rank on its profitability growth measure (instead of the sales target measure) in year $t-1$. Finally, we use the raw value of the RPE profitability measure in year $t-1$. All inferences remain unchanged under the first and second alternative measures. When we use the third alternative measure, the coefficient is significantly negative at the 5 percent level for regular mail, but is insignificant for strategic mail ($p = 0.21$, one-tailed).

Alternative Measures of Status

We also examine the robustness of our findings related to conflict avoidance using an alternative measure of status. Specifically, we use a measure that captures a store manager's status relative to other store managers in the same region, rather than relative to the regional director. This measure is constructed by creating an indicator variable that equals 1 for store managers whose job grade level corresponds to grade 4, and 0 for store managers whose job grade level corresponds to grade 5. All inferences remain unchanged when we re-estimate Equations (1) and (2) using this alternative measure of status.

Potential Scale Effects

To ensure that the results are robust to potential scale effects, we deflate the dependent variables *ADJUST*, *ATTAIN1*, and *ATTAIN2* by target sales and re-estimate all regression equa-

tions. The re-estimated results (not tabulated) are similar to those reported in Tables 4 and 5.

Additional Controls

We also perform additional tests to control for two alternative explanations for our results. First, more experienced regional directors may have greater ability to unravel the slack-building behavior of store managers with private information. To test for this possibility, we create the indicator variable *EXPER*, which takes the value of 1 for directors who are internally promoted and 0 for directors hired from outside of POST, and interact this variable with *N_SUB*, one of our information asymmetry variables. The results (not tabulated) show that the interaction term between *EXPER* and *N_SUB* is insignificant for regular mail ($p = 0.39$, one-tailed) and only marginally significantly negative for strategic mail ($p = 0.08$, one-tailed). All other inferences remain unchanged.

Second, we examine whether unknown individual store manager traits that are correlated with store characteristics drive our results. We analyze this by calculating the average store observation for all variables and using these newly created variables to estimate Equations (1) and (2). Because the mean store manager tenure is less than one year,[25] averaging yearly observations for a given store eliminates any potential effect of store manager traits on target adjustments. The results are robust to averaging yearly observations for a given store.

VII. CONCLUDING REMARKS

Using 2001 to 2003 annual performance evaluation data on 103 postal stores in Korea, we analyze how regional directors use their discretion to allocate regional sales targets set by corporate headquarters to individual stores in the region. Our findings are consistent with predictions based on agency and behavioral theories that regional directors strategically use discretion in the target-setting process to manage compensation risk and address fairness concerns. The results are also consistent with our prediction based on research in psychology that regional directors provide easier targets to store managers with relatively higher hierarchical status to reduce confrontation costs.

Our study has several important implications. First, our results suggest that supervisor discretion is used strategically to improve incentive contracting. In our setting, some stores' initial sales targets based on the regional targets set by corporate headquarters may impose excessive risk because of higher store-specific levels of environmental uncertainty. We find evidence that supervisors manage this risk by using their discretion in the target-setting process to assign more attainable targets to stores facing higher levels of store-specific risk. These adjustments effectively provide a risk premium for bearing more risk.

Our study also suggests that supervisors strategically use discretion in the target-setting process to deal with fairness concerns. Importantly, this study is the first to provide field data evidence consistent with supervisors exercising discretion in the target-setting process to counter other elements of the incentive contract that may be considered unfair. This strategic use of discretion to harmonize the elements of the compensation contract underscores the importance of studying the compensation system as a whole. While the majority of studies that examine compensation contracts focus on one or two elements in isolation, our results suggest that supervisors view different elements of the compensation contract as interrelated. Thus, when a manager has no influence over a certain dimension of the compensation contract but has discretion over others, spillover effects are likely to occur (Bol and Smith 2009).

[25] See footnote 13.

Our findings also emphasize the importance of considering supervisors' personal preferences and incentives when analyzing the target-setting process. We provide evidence consistent with supervisors using their discretion in target setting not only to improve compensation contracting, but also to decrease the personal costs involved in managing the performance evaluation process. Our results suggest that supervisors reduce confrontation costs by strategically determining who to disappoint with relatively harder-to-attain targets and who to satisfy with more attainable targets. In particular, supervisors appear to disappoint those subordinates who are less likely to confront them and as a result avoid painful conflicts as much as possible. Viewed from a decision-theory perspective, the supervisors seem to try to maximize their expected utility over both pecuniary and nonpecuniary elements.

More generally, our study illustrates that discretion is relevant at more than just the performance evaluation stage. Most studies that examine the role of supervisor bias focus on supervisor discretion in the evaluation phase of the compensation process. In contrast, our study shows that biasing behavior may also take place when agreeing on the specifics of the incentive system. Future research could examine what drives principals to provide supervisors with discretion in target setting in certain situations, while allowing discretion in evaluation in other settings. Similarly, in a setting where supervisors are allowed multiple forms of discretion, future research could analyze when supervisors prefer to use their discretion in the target-setting process compared to their discretion in *ex post* evaluation.

This study is subject to several limitations, the most important of which is the fact that the observations come from a single firm, limiting the extent to which we can generalize the results. Another important caveat is the lack of access to the store managers' sales target proposals, because these proposals likely impact regional directors' use of discretion in setting targets. While we attempt to capture the effect of store managers' proposals by controlling for information asymmetry, as store managers' informational advantage is expected to be reflected in their target proposals, we are unable to capture this dimension directly because of data limitations. Future research could collect more complete information on the participative element of the target-setting process and investigate how store managers' proposals affect supervisors' use of discretion in the target-setting process.

APPENDIX

To develop a reliable prediction model for future sales, we use a comprehensive set of independent variables that are likely to explain much of the variation in store-level sales. Our main explanatory variable is prior-year sales because it is likely the key input in estimating this year's sales. In addition, we include the variables that POST used in their cluster analysis to create the RPE reference groups (Matsumura and Shin 2006).[26] POST believes that these variables are predictive of store profitability; because profitability and sales are highly correlated, it is reasonable to assume that these variables are also predictive of sales. We use all explanatory variables measured in year $t-1$ under the assumption that the regional directors will make their own forecasts based on the information available at the beginning of year t. Because there are two sales targets, regular and strategic, we estimate the following cross-sectional regressions using the dependent variables *ACT_SALES_REG* and *ACT_SALES_STRA*:

$$\begin{aligned} ACT\_SALES\_REG_{i,t} = {} & \alpha_0 + \alpha_1 ACT\_SALES\_REG_{i,t-1} + \alpha_2 SERVICE\_AREA_{i,t-1} + \alpha_3 POP_{i,t-1} \\ & + \alpha_4 N\_BANK_{i,t-1} + \alpha_5 METRO_{i,t-1} + \alpha_6 NEW\_MGR_{i,t-1} + \alpha_7 DIST_{i,t-1} \\ & + \alpha_8 SKILL_{i,t-1} + YEAR\_INDICATORS + \varepsilon_{i,t} \end{aligned} \quad (3)$$

[26] Among the variables POST used in their cluster analysis, we do not use the number of employees (*N_EMP*) in our prediction model because that variable is used as a predictor when estimating columns (1) and (2) of Table 5.

$$ACT\_SALES\_STRA_{i,t} = \alpha_0 + \alpha_1 ACT\_SALES\_STRA_{i,t-1} + \alpha_2 SERVICE\_AREA_{i,t-1} + \alpha_3 POP_{i,t-1} + \alpha_4 N\_BANK_{i,t-1} + \alpha_5 METRO_{i,t-1} + \alpha_6 NEW\_MGR_{i,t-1} + \alpha_7 DIST_{i,t-1} + \alpha_8 SKILL_{i,t-1} + YEAR\_INDICATORS + \varepsilon_{i,t} \quad (4)$$

where:

ACT_SALES_REG = store-level actual sales from regular mail (in billion Korean Won);
ACT_SALES_STRA = store-level actual sales from strategic mail (in billion Korean Won);
SERVICE_AREA = area that the store is responsible for (in hundreds of square kilometers);
POP = natural log of the population in the service area;
N_BANK = number of financial institutions in the service area (in hundreds);
METRO = 1, 2, 3, and 4 for rural area, small-sized city, middle-sized city, and metropolitan area, respectively;
NEW_MGR = indicator variable that is equal to 1 if the new manager is appointed in the given year, and 0 otherwise;
DIST = average daily delivery distance per mail carrier (in hundred kilometers);
SKILL = natural log of average labor cost per employee.

We use the store-year-level predicted values of sales to proxy for regional directors' sales estimates. The adjusted R^2s from estimating Equations (3) and (4) are 94 percent and 70 percent, respectively, suggesting that a substantial part of the variation in sales can be explained by the set of store-specific information available to the regional directors at the beginning of year *t*.

REFERENCES

Akerlof, G. A., and J. L. Yellen. 1988. Fairness and unemployment. *The American Economic Review* 78 (2): 44–49.

Anderson, C., and J. L. Berdahl. 2002. The experience of power: Examining the effects of power on approach and inhibition tendencies. *Journal of Personality and Social Psychology* 83 (6): 1362–1377.

———, O. P. John, D. Keltner, and A. M. Kring. 2001. Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology* 81 (1): 116–132.

Anderson, S. W., H. C. Dekker, and K. L. Sedatole. 2010. An empirical examination of goals and performance-to-goal following the introduction of an incentive bonus plan with participative goal setting. *Management Science* 56 (1): 90–109.

Baiman, S., and J. H. Evans, III. 1983. Pre-decision information and participative management control systems. *Journal of Accounting Research* 21 (2): 371–395.

Banker, R., and R. Mashruwala. 2007. The moderating role of competition in the relationship between nonfinancial measures and financial performance. *Contemporary Accounting Research* 24 (3): 763–793.

Berry, T., J. Bizjak, M. Lemmon, and L. Naveen. 2006. Organizational complexity and CEO labor markets: Evidence from diversified firms. *Journal of Corporate Finance* 12 (4): 797–817.

Blanchard, G. A., C. W. Chow, and E. Noreen. 1986. Information asymmetry, incentive schemes, and information biasing: The case of hospital budgeting under rate regulation. *The Accounting Review* 61 (1): 1–15.

Bol, J. C. 2009. The determinants and performance effects of supervisor bias. Working paper, University of Illinois at Urbana–Champaign.

———, and S. D. Smith. 2009. The effects of objective performance measures on subjective performance evaluation. Working paper, University of Illinois at Urbana–Champaign.

Bruggen, A., and J. L. Luft. 2009. Capital rationing, competition, and misrepresentation in budget forecasts. Working paper, Maastricht University and Michigan State University.

Bushman, R., R. Indjejikian, and A. Smith. 1995. Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies. *Journal of Accounting Research* 33 (Supplement): 101–128.

———, Q. Chen, E. Engel, and A. Smith. 2004. Financial accounting information, organizational complexity, and corporate governance systems. *Journal of Accounting and Economics* 37 (2): 167–201.

Christensen, J. 1982. The determination of performance standards and participation. *Journal of Accounting Research* 20 (Autumn): 589–603.

Cichello, M. S., C. E. Fee, C. J. Hadlock, and R. Sonti. 2009. Promotions, turnover, and performance evaluation: Evidence from the careers division managers. *The Accounting Review* 84 (4): 1119–1143.

Cohen-Charash, Y., and P. E. Spector. 2001. The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes* 86 (2): 278–321.

Colquitt, J. A., D. E. Conlon, M. J. Wesson, C. O. L. H. Porter, and K. Y. Ng. 2001. Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *The Journal of Applied Psychology* 86 (3): 425–445.

Cropanzano, R., and R. Folger. 1989. Referent cognitions and task decision autonomy: Beyond equity theory. *The Journal of Applied Psychology* 74 (2): 293–299.

De With, E., and A. Dijkman. 2008. Budgeting practices of listed companies in the Netherlands. *Management Accounting Quarterly* 10 (1): 26–36.

Dunk, A. S. 1993. The effect of budget emphasis and information asymmetry on the relation between budgetary participation and slack. *The Accounting Review* 68 (2): 400–410.

———, and H. Nouri. 1998. Antecedents of budgetary slack: A literature review and synthesis. *Journal of Accounting Literature* 17: 72–96.

Evans, J. H., R. L. Hannan, R. Krishnan, and D. V. Moser. 2001. Honesty in managerial reporting. *The Accounting Review* 76 (4): 537–559.

Govindarajan, V. 1986. Impact of participation in the budgetary process on managerial attitudes and performance: Universalistic and contingency perspectives. *Decision Sciences* 17 (4): 496–516.

Holmström, B. 1979. Moral hazard and observability. *The Bell Journal of Economics* 10 (1): 74–91.

Holthausen, R., D. Larcker, and R. Sloan. 1995. Business unit innovation and the structure of executive compensation. *Journal of Accounting and Economics* 19 (2): 279–313.

Indjejikian, R. J., and M. Matejka. 2006. Organizational slack in decentralized firms: The role of business unit controllers. *The Accounting Review* 81 (4): 849–872.

———, and D. Nanda. 2002. Executive target bonus and what they imply about performance standards. *The Accounting Review* 77 (4): 793–819.

Ittner, C. D., and D. F. Larcker. 2001. Assessing empirical research in managerial accounting: A value-based management perspective. *Journal of Accounting and Economics* 32 (1–3): 349–410.

Keating, S. 1997. Determinants of divisional performance evaluation practice. *Journal of Accounting and Economics* 24 (3): 243–273.

Konovsky, M. A. 2000. Understanding procedural justice and its impact on business organizations. *Journal of Management* 26 (3): 489–511.

Lambert, R. A. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32 (1–3): 3–87.

Libby, T. 2001. Referent cognitions and budgetary fairness: A research note. *Journal of Management Accounting Research* 13 (1): 91–105.

Lindquist, T. M. 1995. Fairness as an antecedent to participative budgeting: Examining the effects of distributive justice, procedural justice and referent cognitions on satisfaction and performance. *Journal of Management Accounting Research* 7: 122–147.

Locke, E. A., and G. P. Latham. 1990. *A Theory of Goal-Setting and Task Performance*. Englewood Cliffs, NJ: Prentice Hall.

Lukka, K. 1988. Budgetary biasing in organizations: Theoretical framework and empirical evidence. *Accounting, Organizations and Society* 13 (3): 281–301.

Matsumura, E. M., and J. Y. Shin. 2006. An empirical analysis of an incentive plan with relative performance

measures: Evidence from a postal service. *The Accounting Review* 81 (3): 533–566.

Merchant, K. A. 1985. Budgeting and the propensity to create budgetary slack. *Accounting, Organizations and Society* 10 (2): 201–210.

———, and J.-F. Manzoni. 1989. The achievability of budget targets in profit centers: A field study. *The Accounting Review* 64 (3): 539–558.

Murphy, K. J. 2000. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30 (3): 245–278.

Murphy, K. R., and J. Cleveland. 1995. *Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives*. Thousand Oaks, CA: Sage Publications.

Nagar, V. 2002. Delegation and incentive compensation. *The Accounting Review* 77 (2): 379–395.

Phillips, E., and R. Cheston. 1979. Conflict resolution: What works? *California Management Review* 21 (4): 76–83.

Rahim, M. A. 1986. Referent roles and styles of handling interpersonal conflict. *The Journal of Social Psychology* 126 (1): 79–86.

Rogers, W. 1993. *Regression Standard Errors in Clustered Samples*. Stata Technical Bulletin Reprints, Vol. 3: 83–94. College Station, TX: Stats Press.

Schiff, M., and A. Y. Lewin. 1968. Where traditional budgeting fails. *Financial Executive* 36 (5): 51–62.

———, and ———. 1970. The impact of people on budgets. *The Accounting Review* 45 (2): 259–268.

Schoute, M., and E. Wiersma. 2007. The impact of budgetary slack on budget use. Working paper, Vrije Universiteit Amsterdam.

Shastri, K., and D. E. Stout. 2008. Budgeting: Perspectives from the real world. *Management Accounting Quarterly* 10 (1): 18–25.

Thye, S. R. 2000. A status value theory of power in exchange relations. *American Sociological Review* 65 (3): 407–432.

———, D. Willer, and B. Markovsky. 2006. From status to power: New models at the intersection of two theories. *Social Forces* 84 (3): 1471–1495.

Umapathy, S. 1987. *Current Budgeting Practices in U.S. Industry*. New York, NY: Quorum Books.

Young, S. M. 1985. Participative budgeting: The effects of risk aversion and asymmetric information on budgetary slack. *Journal of Accounting Research* 23 (2): 829–842.

Yukl, G., and C. M. Falbe. 1990. Influence tactics and objectives in upward, downward, and lateral influence attempts. *The Journal of Applied Psychology* 75 (2): 132–140.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 1887–1919

American Accounting Association
DOI: 10.2308/accr.2010.85.6.1887

Determinants of Hedge Fund Internal Controls and Fees

Gavin Cassar
University of Pennsylvania

Joseph Gerakos
University of Chicago

ABSTRACT: We investigate the determinants of hedge fund internal controls and their association with the fees that funds charge investors. Hedge funds are subject to minimal regulation. Hence, hedge fund managers voluntarily implement internal controls, and managers and investors freely contract on fees. We find that internal controls are stronger in funds with higher potential agency costs. Further, internal controls are stronger in funds domiciled in jurisdictions that provide investors with limited legal redress for fraud and financial misstatements. Short selling funds, however, are more likely to protect information about their investment positions by implementing weaker internal controls. With respect to fees, we find that the percentage of positive profits that the manager receives increases in the strength of the fund's internal controls. Finally, removing the manager from setting and reporting the fund's official net asset value, along with reputational incentives and monitoring by leverage providers, are all associated with lower likelihoods of future regulatory investigations of fraud and/or financial misstatement.

Keywords: *hedge funds; internal controls; investor fees; restatements.*

Data Availability: *The data used in this study are obtained from TASS and through agreement with The Hedge Fund Due Diligence Group at Analytical Research (HedgeFundDueDiligence.com). This agreement prohibits the authors from making the data available to other parties.*

The authors thank The Hedge Fund Due Diligence Group at Analytical Research (HedgeFundDueDiligence.com) for providing the data used in the project. The authors also thank two anonymous referees, Steve Kachelmeier (editor), Hollis Ashbaugh-Skaife, Ralph DiMeo, Wayne Guay, Chris Ittner, Robert Krause, Doug Skinner, Ro Verrecchia, and seminar participants from the 2009 AAA Annual Meeting, Florida State University, Harvard Business School's 2009 IMO Conference, Maastricht University, McGill University, Northwestern University, Tilburg University, University of Washington, and the Wharton Applied Economics Workshop for their comments. This project benefited from the research assistance of Amy Chang, Tanoya Chatterjee, Jessica Leung, and Nicholas Smith-Wang. The financial support provided by The Wharton School and The Booth School's Initiative on Global Markets is greatly appreciated.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: January 2009
Accepted: February 2010
Published Online: November 2010

I. INTRODUCTION

Hedge funds encompass a diverse range of privately managed investment vehicles that are exempt from a broad range of federal acts regulating investment vehicles. They are typically not exchange traded and not registered with the Securities and Exchange Commission (SEC). Hence, hedge funds are in general exempt from securities regulations that dictate internal controls that managers must implement and maintain, fees that managers can charge investors, and disclosures that fund managers must make to investors. Therefore, hedge funds are opaque investment vehicles that expose investors to the risk of substantial losses arising from fraud and/or financial misstatements. Hedge funds thereby provide a setting to examine the extent to which fund managers voluntarily implement internal controls to limit agency costs arising from fraud and/or financial misstatements, and the extent to which investors value such internal controls when determining the fees that they are willing to pay fund managers.

In this study, we examine the determinants of hedge fund internal controls and the association between internal controls and the fees that fund managers charge investors. We use a broad definition of internal controls that encompasses mechanisms designed to decrease the likelihood of fraud, and to increase the accuracy of asset valuations and performance disclosures made to investors. Some examples include the independent pricing of investment positions, signature protocols for transferring funds from bank and prime brokerage accounts, and the use of reputable service providers, such as auditors and administrators. The risks that hedge fund internal controls monitor and reduce could be more important than financial risks in determining hedge fund performance (Lo 2001, 30–31; Kundro and Feffer 2003).

To investigate the use of internal controls and their association with fees, we use a proprietary database of due diligence reports prepared by The Hedge Fund Due Diligence Group at Analytical Research (HedgeFundDueDiligence.com). These reports contain an extensive array of details regarding fund characteristics and internal controls, investment contract terms and provisions, investment style and portfolio characteristics, and fund and manager backgrounds. Investors commissioned these due diligence reports to evaluate whether to invest in the funds. Therefore, our sample represents a set of hedge funds that were actively seeking investors and for which investors initiated due diligence to evaluate the backgrounds, contract terms, and internal controls of the funds.

We find substantial variation in internal controls. Funds domiciled in the Caribbean are more likely to incorporate stricter signature authority for the transfer of funds, implement pricing mechanisms that involve external oversight, and employ more reputable outside service providers. These findings are consistent with investors demanding greater internal controls for funds domiciled in jurisdictions that provide investors with limited legal redress for fraud and financial misstatements. Furthermore, and also consistent with managers implementing internal controls to reduce agency conflicts, we find that younger funds are more likely to use more reputable external administrators and levered funds are more likely to implement more stringent external oversight of pricing and use more reputable auditors. Short selling funds, however, are more likely to protect proprietary information by implementing weaker internal controls that limit external oversight of investment position pricing.

Next, we argue that when considering a hedge fund investment, investors estimate potential agency costs arising from fraud and financial misstatements, and their expectations of these agency costs decrease in the quality of a fund's internal controls. Consequently, we posit that managers of funds with more stringent internal controls can charge higher fees, all else equal. Consistent with this argument, we observe that internal controls that reduce managers' opportunities to manipulate reported performance and/or commit fraud are positively associated with the

percentage of investment profits received by the manager. Therefore, investors appear to mitigate moral hazard costs arising from fraud and/or manipulated reported returns when managers have greater discretion.

We supplement the above analysis of fees by examining the association between restatements and fees. Relevant to our setting, restatements are predetermined at the point of contracting. We find that managers of funds that have restated prior performance receive a significantly lower percentage of assets under management for managing the funds, providing further support for our prediction that investors protect against the risk of future misstatements by paying lower fees.

Finally, we investigate whether internal controls are associated with future regulatory investigations of fraud and/or financial misstatements on the part of the fund and its managers. Such investigations can be considered an extreme form of poor performance because they are typically associated with fund liquidations and deeply discounted investor redemptions. We find that excluding the manager from setting and reporting the fund's official net asset value (NAV) to investors reduces the probability of such investigations by over one-third. Moreover, we find that proxies for reputational incentives and proxies for explicit monitoring and screening by leverage providers are associated with lower likelihoods of future regulatory investigations of fraud and/or financial misstatements.

We contribute to the internal control and hedge fund literatures. First, this study differs from previous research on internal controls by examining the supply and demand of actual internal controls determined within a market as opposed to the disclosure of internal control weaknesses.[1] Further, research is generally based on public firms subject to securities regulations that limit the risks that investors face from weak internal controls. In contrast to public companies, internal control failures can be more important than financial risks in determining hedge fund performance (Lo 2001, 30–31). Moreover, research that investigates internal control choice in non-regulated settings focuses primarily on auditor choice (e.g., Chow 1982; Watts and Zimmerman 1983; Blackwell et al. 1998; Fortin and Pittman 2007). We also contribute by investigating a broad range of internal controls in a voluntary setting with the potential for substantial agency costs. Our empirical results provide several insights. We find in a voluntary setting that fund reputation substitutes for explicit internal controls and that legal domicile influences the choice of internal controls employed. We also find that weaker internal controls are associated with a lower percentage of investment profits received by the manager. Further, manager involvement in setting and reporting the fund's official NAV increases the likelihood of future regulatory investigation for fraud and financial misstatement. But, we find no support for higher quality auditors reducing the likelihood of future regulatory investigations.

Second, we extend the literature on hedge funds by examining their internal operations. Research on hedge funds generally concentrates on investment returns (see Lo [2008] for a review). We extend this research by directly observing the internal controls employed by funds to reduce agency costs, and how these controls are associated with fees. Moreover, we demonstrate that weak internal controls are associated with future regulatory investigations of fraud and financial misstatements and that reputational incentives and leverage providers reduce the likelihood that managers carry out actions that lead to such regulatory investigations.

Finally, we contribute to the recent debate over the regulation of hedge funds. In general, the SEC regards internal controls as a critical element of investor protection, and it recently increased its regulatory focus on hedge funds, proposing regulations that include mandatory disclosures and other internal controls (Smith 2006a, 2006b; Oesterle 2006). Hedge fund advocacy groups re-

[1] Some examples are Ashbaugh-Skaife et al. (2007), Ashbaugh-Skaife et al. (2009), Doyle et al. (2007), Ge and McVay (2005), Hammersley et al. (2008), Kim et al. (2008), and Ogneva et al. (2007).

sponded to these proposals by suggesting that funds follow "best practice" industry standards that give consideration to the particular characteristics and circumstances of each fund (Managed Funds Association [MFA] 2005). Consistent with the advocacy groups' proposals, we find that funds systematically vary their internal controls to address potential agency costs. Furthermore, we show that internal controls are positively associated with investor fees, suggesting that investors evaluate and price the risks arising from internal control failures.

II. HEDGE FUND SETTING

Hedge funds are managed investment vehicles. They are often privately held, generally comprised of wealthy individuals and institutional investors, and typically organized in the U.S. as a limited partnership or offshore as a corporation (Fung and Hsieh 1999).[2] There has been substantial growth in the hedge fund industry, both in the number of funds and in assets under management (Brown et al. 1999; Fung and Hsieh 1999; Lo 2008). As of the first quarter of 2008, hedge funds held more than $2.8 trillion in assets under management (HedgeFund.net [HFN] 2008). Although hedge funds have grown tremendously and are under intense scrutiny regarding their operations and potential contribution to systemic risk, they are opaque, and therefore little is known about how they operate (Lo 2001; Chan et al. 2006).

Unlike other investment vehicles, hedge funds are structured to be exempt from the public offering requirements of the Securities Act of 1933, the periodic reporting obligations of the Securities Exchange Act of 1934, and the registration requirements of the Investment Company Act of 1940 (Oesterle 2006).[3] For example, Section 5 of the Securities Act of 1933 requires registration of the sale of securities unless the issue qualifies for a "private placement" exemption.[4] Rule 506 of Regulation D provides exemption for hedge funds if they restrict their offerings to only accredited investors, namely investment companies, or individuals with more than $1 million or income exceeding $200,000 in the two most recent years or joint income with a spouse of over $300,000, and up to 35 other U.S. purchasers.[5] Additionally, to qualify for exemption, all non-accredited investors must be determined by the fund to be sophisticated, with knowledge and experience to evaluate the prospective investment. Therefore, a more descriptive definition of a hedge fund is an investment fund exempt from a list of specific federal acts regulating investment vehicles (Oesterle 2006).[6] To ensure exemption from SEC regulation, hedge funds cannot undertake any form of general solicitation or advertising for their services or sale of securities to the general public, and, therefore, must solicit investments through private placements only to those who are sophisticated enough to evaluate the investment and have sufficient wealth to bear the risk of the investment (American Bar Association [ABA] 2005, 214–215; Lhabitant 2008).[7]

This minimal regulatory environment provides hedge funds substantially more flexibility than regulated investment vehicles, such as mutual funds. For example, hedge funds have greater discretion regarding valuation and reporting of their investments (SEC 2003; McVea 2008). Unlike other investment vehicles that are registered under the Investment Company Act of 1940,

---

[2] Although the early hedge funds generally held offsetting long and short positions (leading to the title "hedge funds"), hedge funds are not limited to strict long-short positions.

[3] For a more detailed discussion of the regulation of hedge funds see Fung and Hsieh (1999, 314–317), ABA (2005), and Lhabitant (2008, 37–84).

[4] Although issuers of private placements are required to file a Form D with the SEC within 15 days after the first sale, this form contains very little detailed information about the issuance (Lhabitant 2008, 41).

[5] Regulation S provides detailed exemptions related to issuances to non-U.S. investors.

[6] Providing a definition of a hedge fund based on operating characteristics is problematic given its wide variation in investment asset class, sector, duration, region, and organizational structure (ABA 2005, 1; Lhabitant 2008, 1).

[7] Violation of this solicitation rule could evoke registration of the sale of securities under the Securities Act of 1933, registration as a mutual fund under the Investment Company Act of 1940, and registration as an investment advisor under the Investment Advisers Act of 1940.

hedge funds have greater discretion to use leverage to finance their investment positions and can undertake substantial short selling (Lhabitant 2008, 46).[8] In addition, hedge funds can charge fees based on performance, whereas other investment vehicles, such as mutual funds, are restricted to fees based solely on assets under management (Fung and Hsieh 1999; ABA 2005, 331–334).

When evaluating an investment in a hedge fund, investment advisors or accredited investors solicit information about the fund, with the investment terms provided in an offering circular or "private placement memorandum" (PPM). The PPM, and the subsequent executable limited partnership agreements and subscription agreements, lay out the fund's operations and the investor's contractual rights: the fund's investment strategy, the fees agreed to be paid, the terms under which the investor can invest and withdraw funds, the investor's ability to monitor the fund, the manner and frequency in which the fund will estimate and report performance, and the investor's remedy rights in the case of a dispute (ABA 2005, 96–98). Because hedge funds are substantively exempt from securities regulations, the contractual terms laid out in the PPM and subsequent executable agreements represent the primary mechanisms in place to protect the investor's investment.

III. PREDICTIONS

Hedge funds are opaque due to the limited mandatory disclosure requirements and the proprietary nature of their operations. Consequently, there is the potential for substantial agency costs and information asymmetries between investors and fund managers. Moreover, hedge funds often engage in complex trading strategies that involve large numbers of securities and/or trades, thereby increasing the potential for fraud on the part of employees. Therefore, this setting presents strong incentives to implement internal controls that increase the monitoring of managers, reduce the likelihood of fraud, and increase the accuracy of asset valuations and performance disclosures made to investors.[9] Consistent with the importance of hedge fund internal controls, Kundro and Feffer (2003) find that the breakdown of internal systems was a major determinant of failure for more than 50 percent of the 100 failed funds they investigated.

The fund manager is the residual claimant on the fund. By implementing internal controls to reduce the agency costs between investors and the fund, managers can increase the value of their residual claim if the benefits of implementation, such as obtaining greater fees, are greater than the costs of implementing and maintaining internal controls. Whereas there are likely to be benefits to having particular internal controls in place, not all hedge fund managers will invest or commit to using internal controls, as they are costly to use and, in many cases, have a large fixed component (SEC 2003). For example, external oversights can inhibit the manager's ability to quickly respond to trading opportunities. Further, there are explicit costs to retain external service providers for operational oversight and support, and the use of external service providers requires the manager to disclose to the service providers proprietary information about investment strategies and positions.

We predict that *ex ante* agency costs are higher for younger funds that lack valuable reputations, offshore funds that are subject to a weaker legal system, and unlevered funds that are not subject to creditor monitoring.[10] Managers of established, higher quality funds have valuable reputations and potentially superior investing technologies, implying that they hold more valuable

[8] For example, open-end investment companies are restricted to bank-sourced leverage only, with this amount capped at 300 percent of the asset coverage (Section 18(f)(1) Investment Company Act of 1940).

[9] Note that even in the absence of agency problems, internal controls can increase the precision of asset values disclosed to investors and, therefore, reduce uncertainty about the cash flows that investors expect to receive.

[10] We present our predictions and empirics using the fund as the unit of analysis, because any one manager can manage both levered and unlevered funds, and both offshore and onshore funds. Further, prior literature exclusively employs fund-level variables in empirical analyses. While both fund and manager characteristics could potentially affect internal

residual claims from operating their funds than do managers of lower quality funds. Therefore, managers of higher quality funds have lower incentives to commit fraud and greater incentives to provide investors with more accurate performance reports (Klein and Leffler 1981), reducing the benefit of explicit internal controls. Fund quality can be signaled by past investment performance and capital flows (Fung and Hsieh 1997). For new funds, there is greater uncertainty about the quality of the manager and the likelihood that the manager will commit fraud or misstate performance. This uncertainty should increase the benefit to new funds from enacting more stringent internal controls and using reputable outside service providers. For example, Mansi et al. (2004) find that auditor quality and tenure are negatively related with the cost of debt financing, especially for firms with riskier grades of debt. Consequently, *ceteris paribus*, we predict a negative association between internal controls and fund age.[11,12]

Although onshore and offshore funds are generally exempt from U.S. securities regulations, investors in onshore funds can use the U.S. legal system to redress fraud and financial misstatements.[13] Although many offshore funds are domiciled in countries that regulate investment vehicles and provide investors with legal redress, U.S. legal contract enforcement mechanisms are generally considered superior. Most offshore hedge funds are domiciled in Caribbean islands that have historically been associated with secret bank accounts and money laundering (Suss et al. 2002). Fund managers who operate in a regulatory environment that allows secret bank accounts and enables money laundering can find that the lax banking environment makes it easier to commit fraud, and, conditional on committing fraud, easier to hide and protect such fraudulent gains from defrauded investors (for a discussion, see Blum et al. 1998).

There could be systematic variation in the types of investors that invest in onshore versus offshore funds.[14] In particular, when U.S. tax-exempt entities, such as pension funds and charitable trusts, invest in offshore funds, they can avoid recognizing unrelated business taxable income arising from investments in leveraged portfolios (ABA 2005, 360; Lhabitant 2008, 88). Nevertheless, it is not clear why, holding other investor characteristics constant, U.S. tax-exempt investors would have monitoring incentives that differ significantly from those of other investors, which suggests that this potential difference has limited implications for hedge fund internal controls. Therefore, *ceteris paribus*, we predict that funds domiciled offshore will have better internal controls because investors have reduced legal recourse to redress fraud and financial misstatements.

---

control choice, inclusion of both manager and fund characteristics results in multicollinearity arising from the fact that many managers run only one fund in our sample. In our "Robustness" section we examine the choice of using either the fund or manager as the unit of analysis, and including manager characteristics in our tests.

[11] Smaller funds can be considered to be of lower quality because they have had less capital inflows. Therefore, they are likely to have greater benefits from implementing internal controls than larger funds. However, larger funds have greater incentives to implement internal controls because the magnitude of potential agency costs increases in fund size. For example, research has consistently shown a positive association between better internal controls and organization size (Chow 1982; Doyle et al. 2007). Therefore, *ex ante*, it is unclear what the association between fund size and better internal controls will be.

[12] Note that this prediction is based on the assumption either that internal control technology has improved over time and older funds find it too costly to implement newer technologies, or that over time funds receive decreasing benefits from internal controls and therefore unravel internal controls.

[13] If the manager is onshore but the fund is offshore, investors can sue the fund manager through the U.S. legal system. Nevertheless, any judgments against the fund would be executed through the legal system of the fund's offshore domicile.

[14] Although the Taxpayer Relief of 1997 eliminated the historical primary tax rationale for offshore hedge fund investing, namely the avoidance of U.S. taxation of non-U.S. investors on their capital gains, non-U.S. investors could still prefer offshore funds to limit their exposure and retain their anonymity from U.S. regulatory authorities, such as the IRS (ABA 2005, 98–99 and 360; Lhabitant 2008, 88). Like previous hedge fund research, we are unable to identify specific investors in our sample funds. The extent to which we cannot control for investor is a limitation of the research.

Some funds use leverage to increase their invested assets. Funds generally obtain leverage from prime brokers who provide managers with trade execution and custodial services (Lhabitant 2008). Leverage creates another stakeholder in the fund's performance and operations. Therefore, leverage providers are likely to demand that funds implement strong internal controls as precondition for providing credit, leading to a positive relation between leverage and internal controls. However, to the extent that the prime broker has an incentive and the ability to monitor the fund manager because it typically executes the fund's trades, leverage could dampen this hypothesized association, given potential substitutability between formal internal control mechanisms and monitoring on the part of debt holders.

Finally, investment style can affect the costs associated with a fund's internal control choices. It is generally accepted that a hedge fund investment is essentially a bet on the fund manager's proprietary investment strategies (Edwards and Caglayan 2001; Lo 2008). Hence, a fund can incur substantial costs if outsiders acquire proprietary information about the fund's investment positions.[15] These costs are particularly relevant for funds that employ a short bias investment style. Outsiders who learn a fund's short positions can profit because the fund eventually has to unwind its short positions, and, therefore, informed outsiders can trade against the fund when fund has to cover its short positions (Dechow et al. 2001). If short bias funds take actions to protect proprietary information about investment positions, then they should limit the use of objective external verification of fund operations, such as the valuation of fund assets.[16] Consequently, *ceteris paribus*, we predict a negative association between short selling and internal controls that involve outside service providers.

In summary, we predict that younger funds and offshore funds are more likely to adopt stronger internal controls and reporting mechanisms, as these funds are riskier along the dimensions of fraud and financial misstatements and, therefore, internal controls should allow managers of these funds to charge higher fees. In addition, we predict that leverage affects managers' internal control decisions and that short bias funds are less likely to use external parties to value investment positions.

IV. SAMPLE

To investigate the determinants of internal controls and their association with fees, we use a database of hedge fund due diligence reports prepared by HFDD. The reports were commissioned by investors evaluating whether to invest in the funds.[17,18] Therefore, the sample represents a set of hedge funds in which investors were actively seeking to invest capital. HFDD specializes exclusively in hedge fund due diligence, and it obtained the information included in these reports

---

[15] Note that hedge funds are required under Section 13(d) and Section 13(g) of the Securities Act of 1933 to report long positions relevant to corporate control and transfer, namely those with more than 5 percent of a class of equity security registered under Section 12 of the Securities Exchange Act 1934. Also, hedge funds with discretion over $100 million in assets are required to disclose their long positions on a quarterly basis. However, these requirements are of limited proprietary risk given the disclosures are aggregated, delayed, and restricted to long positions only (Lhabitant 2008, 44).

[16] Although service providers can and do implement mechanisms to protect the release and use of clients' proprietary information, such systems are not perfect. For example, Geczy and Yan (2006) find that such mechanisms do not completely prevent the transmission of proprietary information between brokers and market makers, and Ivashina and Sun (2007) find that institutional investors trade on proprietary information that they receive from participating in loan syndicates.

[17] All HFDD investigations were initiated and paid for by potential investors. No investigations were commissioned and/or paid for by the investigated hedge funds.

[18] HFDD is not the only organization that carries out hedge fund due diligence. Due diligence is performed by other third-party providers, and in-house by many fund-of-funds. Like other due diligence providers, HFDD clients are primarily major institutional investors and high net worth individuals who use their services on a repeated basis. Instead of undertaking due diligence, investors can invest in funds of hedge funds that specialize in evaluating and monitoring fund managers (Brown et al. 2008).

from several sources including on-site visits and interviews with key staff, discussions with service providers (auditors and administrators), reviews of offering memorandums, examinations of public filings and registrations, media scans, background checks of fund managers and staff, reviews of financial statements, and reviews of internal control protocols. Consequently, this database overcomes potential concerns related to commercial hedge fund databases that are based on self-reported fund performance and manager characteristics.[19] The HFDD reports provide extensive detail regarding fund characteristics, internal controls, contract terms, and manager backgrounds.

Table 1 provides the descriptive statistics for the funds and managers. The sample consists of 427 funds run by 358 unique managers investigated from 2003 to 2007. Panel A shows the sample distribution by domicile. Sample funds are domiciled primarily in the Cayman Islands (251 funds), United States (70 funds), British Virgin Islands (50 funds), and Bermuda (34 funds). Panel B provides a summary of fund characteristics. The mean (median) fund has $305 million ($107 million) in assets under management and is, on average, less than three years old (1,020 days) at the time of due diligence. Of our sample funds' managers, 36 percent are domiciled offshore; however, in contrast to manager location, most funds (84 percent) are located offshore. Examining the overlap between offshore managers and offshore funds, more than half (56.8 percent) of the sampled offshore funds are managed by onshore managers; however, no onshore funds are managed by offshore managers. For multivariate analyses, we create two indicator variables to reflect the joint distribution of offshore funds and managers, and capture potential variation from the separation of fund and manager domicile. The first variable is coded as 1 if both the manager and the fund are offshore, and 0 otherwise. The second variable is coded as 1 if the fund is offshore and the manager is onshore, and 0 otherwise.

Table 1 also provides descriptive detail on the funds' investment styles, asset classes, and portfolio characteristics. We observe that 54 percent of the funds use leverage to finance their investments. The due diligence firm provided the fund investment styles, which are based quantitative and qualitative analyses of funds' portfolio holdings and returns. Our sample spans a broad range of invested assets including equity, fixed income, and foreign exchange. With respect to investment style, 20 percent of the funds were identified as having a short selling bias, whereas 36 percent were identified as having a long bias.[20] In addition, there is substantial variation in the number of investment positions typically held in the funds' portfolios: 41 percent hold less than 40 positions, and 3 percent hold 1,000 or more positions. There is also variation in the typical holding period for the investments made by the sample funds, with 13 percent holding investment positions for only days and 32 percent holding positions for longer than a year.

Hedge fund managers generally receive two types of fees. First, they receive an annual management fee calculated as a percentage of assets under management. For the sample, the mean (median) management fee is 1.52 percent (1.50 percent) of assets under management. Second, managers receive a performance fee calculated as a percentage of investment profits. For the sample, the mean (median) performance fee is 19.34 percent (20.00 percent).[21] Finally, 10 percent of our sample funds have restated asset values and performance disclosures that were made to investors. Hedge fund restatements are potentially more serious than restatements of public firms,

[19] One concern with the database is that it consists of funds willing to be subject to due diligence. According to HFDD, such refusals are rare and clients typically do not invest in funds that refuse.

[20] The remaining funds can engage in some short or long selling; however, based on quantitative analyses of trading activities, the relative magnitude of active short/long selling is substantially greater for these funds.

[21] A large proportion of management fees are set at 1.00 percent, 1.50 percent, and 2.00 percent, whereas the majority of performance fees are clustered at 20.0 percent. To address the kinks in the distributions of management fees and performance fees, in robustness tests, we replace fee measures with ordinal measures coded as 0 for below the mode of the distribution, 1 equal to the mode, and 2 for above the mode. The results are invariant to this alternative specification.

TABLE 1
Descriptive Statistics of Hedge Funds

Panel A: Domicile of Funds

Domicile	Funds
Cayman Islands	251
United States	70
British Virgin Islands	50
Bermuda	34
Bahamas	5
Other	14
Ambiguous	3
Total	427

Panel B: Descriptive Statistics for Funds

Variable	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
Fund AUM ($ millions)	304.78	652.38	0.00	40.00	107.00	269.00	6300.00
Ln(Fund AUM)	18.43	1.74	0.00	17.50	18.49	19.41	22.56
Fund Age (Days)	1019.59	976.80	0.00	373.00	700.00	1339.00	4877.00
Ln(Fund Age)	6.39	1.30	0.00	5.92	6.55	7.20	8.49
Management Fee (%)	1.52	0.48	0.00	1.00	1.50	2.00	3.50
Performance Fee (%)	19.34	4.42	0.00	20.00	20.00	20.00	50.00
Fund Offshore	0.84						
Manager Offshore	0.36						
Offshore Mgr. and Fund	0.36						
Fund offshore only	0.48						
Leverage	0.54						
Short Bias	0.20						
Long Bias	0.36						
1000+ Positions	0.03						
200–999 Positions	0.08						
100–199 Positions	0.12						

(continued on next page)

Panel B: Descriptive Statistics for Funds

Variable	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
40–99 Positions	0.35						
1–39 Positions	0.41						
Years	0.32						
Quarters	0.31						
Months	0.15						
Weeks	0.09						
Days	0.13						
Restatement	0.10						

Fund AUM is the assets under management for the fund. Fund Age (Days) is the number days since the fund's inception. Management Fee is the percentage of the fund's assets under management that the manager receives annually for managing the fund. Performance Fee is the percentage of positive profits that the manager receives annually as compensation. Fund Offshore is an indicator variable coded as 1 if the fund is registered offshore, and 0 if the fund is located in the United States. Manager Offshore is an indicator variable coded as 1 if the manager is located offshore, and 0 if the manager is located in the United States. Offshore Mgr. and Fund is an indicator variable coded as 1 if both the fund is registered offshore and the manager is located offshore. Fund offshore only is an indicator variable coded as 1 if the fund is registered offshore and the manager is located in the United States.

Leverage is an indicator variable coded as 1 if the fund uses leverage, and 0 otherwise. Short Bias is an indicator variable coded as 1 if the fund's investment style is weighted toward short positions, and 0 otherwise. Long Bias is an indicator variable coded as 1 if the fund's investment style is weighted toward long positions, and 0 otherwise. 1000+ Positions, 200–999 Positions, 100–199 Positions, 40–99 Positions, and 1–39 Positions are indicator variables for the average number of investment positions in the fund's portfolio. Years, Quarters, Months, Weeks, and Days are indicator variables for the average holding period of an investment position. Restatement is an indicator variable coded as 1 if fund has restated the performance reported to its investors, and 0 otherwise.

because hedge fund restatements involve NAVs that determine fees and redemptions. Overall, the characteristics of the sample are consistent with previous empirical evidence on hedge funds (Aragon 2007).

V. RESULTS

Descriptive Statistics of Internal Controls in Hedge Funds

We report a comprehensive set of internal controls used in hedge funds. The first group consists of the signatures required to make transfers out of the fund's bank and prime brokerage accounts. The use of more than one signature authority and the use of external parties to authorize fund movement and withdrawals provide greater control of the cash held by the fund, thereby decreasing the likelihood of fraud. The second group relates to the mechanisms used to estimate and disclose the fund's performance to investors. This group can be further broken down into three subgroups: who prices the portfolio; what is the source of prices used to value individual asset performance; and who calculates the portfolio's NAV that is reported to investors. The use of external parties and objective sources to value invested assets and the use of an independent party without manager involvement to report and verify the NAV to investors provide the most objective measurement of fund position and performance, thereby reducing the likelihood of fraud and increasing the precision of asset valuations disclosed to investors.

The third group of internal controls consists of the quality and reputation of the fund's service providers, specifically its auditor and administrator. Similar to engagements with operating firms, auditors generally undertake annual audits of hedge funds to ensure that the financial statements furnished to fund investors comply with relevant accounting standards. However, although many auditors review the consistency of how funds value investment positions, they do not typically review the validity of valuations (Lhabitant 2008). Nevertheless, auditors with good reputations have incentives to decrease the likelihood of fraud or financial misstatement on the part of their clients.[22] The role of fund administrators varies substantially across engagements; however, they generally provide "back-office" support, such as performing day-to-day administrative operations, accounting and valuation services, and serving as the interface with investors. In many cases the administrator also calculates the NAV using data provided by either the fund custodian or prime broker (Lhabitant 2008). The practitioner literature strongly advocates against allowing fund managers to perform final valuations and/or to communicate valuations to an administrator, except in highly exceptional circumstances, and even then, such occasions are recommended to be made with auditor approval and fully disclosed to investors (Lhabitant 2008, 100). Similar to auditors, we observe that several administrators serve more than one of our sample hedge funds. Therefore, we assume that administrators with valuable reputations work to decrease the likelihood of fraud or financial misstatement.

Table 2, Panel A reports the descriptive statistics for the internal controls implemented by the funds in the sample. We find that the type and number of signatures required to transfer funds out of the bank or prime broker account vary substantially across our sample. We observe 27.7 percent of the funds require multiple signatures from both internal and external parties, and 24.2 percent require only an external signature. Some funds require only internal signatures, with 24.6 percent requiring at least two internal signatures and 23.5 percent requiring only one internal signature to transfer money out of the fund's bank and prime brokerage accounts.

[22] Note that in contrast with much of the auditing literature, we assume that the reputation as opposed to the size of the auditor and/or administrator determines the quality of services provided (see Weber et al. 2008 for a discussion). Although hedge fund service providers can be sued, the reputational loss due to fraud or misstatement is probably larger than any damages that can be recovered especially for funds and service providers located offshore, especially those located in tax havens.

TABLE 2

Descriptive Statistics of Internal Controls

Panel A: Internal Control Measures by Domicile

Variable	Index Weight	Overall (%)	U.S. (%)	Cayman Islands (%)	British Virgin Islands (%)	Bermuda (%)	Other Offshore (%)
Signatures							
One Internal Signature	0	23.5	49.3	20.9	12.2	17.7	5.3
Two Internal Signatures	1	24.6	27.5	22.5	38.8	20.6	10.5
One External Signature	2	24.2	7.3	24.5	26.5	35.3	52.6
Dual/Triple Entity Signatures	3	27.7	15.9	32.1	22.5	26.5	31.6
Who Prices Portfolio							
Manager Only	0	16.2	37.1	11.2	12.0	20.6	5.3
Manager and Administrator	1	11.7	10.0	13.2	8.0	14.7	5.3
Administrator Only	2	62.9	48.6	65.2	76.0	50.0	84.2
Dual/Triple Entity Pricing	3	9.2	4.3	10.4	4.0	14.7	5.3
Least Objective Source of Prices							
Manager	0	13.4	20.3	12.4	16.0	5.9	10.5
Model	1	12.7	18.8	12.0	6.0	14.7	15.8
Dealer Quote	2	33.2	17.4	36.0	38.0	35.3	42.1
OTC Quotes	3	5.9	2.9	5.6	8.0	8.8	0.0
Exchange Quotes	4	34.8	40.6	34.0	32.0	35.2	31.6
Who Sets NAV							
Manager Involved	0	14.1	40.0	8.8	8.0	14.7	5.3
Manager Not Involved	1	85.9	60.0	91.2	92.0	85.3	94.7
Auditor							
Not Ranked by Alpha Magazine	0	23.0	51.4	17.9	20.0	17.7	5.3
Ranked by Alpha Magazine	1	77.0	48.6	82.1	80.0	82.4	94.7
Administrator							
Not Ranked by Alpha Magazine	0	68.4	91.4	59.0	70.0	79.4	78.9
Ranked by Alpha Magazine	1	31.6	8.6	41.0	30.0	20.6	21.1

(continued on next page)

Panel B: Internal Control Indices

	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
Signatures Index	1.56	1.13	0.00	1.00	2.00	3.00	3.00
Who Prices Index	1.65	0.86	0.00	1.00	2.00	2.00	3.00
Source of Prices Index	2.55	1.42	0.00	1.00	3.00	4.00	4.00
NAV No Manager	0.86	0.35	0.00	1.00	1.00	1.00	1.00
Ranked Auditor	0.77	0.42	0.00	1.00	1.00	1.00	1.00
Ranked Administrator	0.32	0.47	0.00	1.00	1.00	1.00	1.00
Internal Control Index	7.52	2.78	0.00	6.00	8.00	10.00	12.00

Signatures refer to the signatures required to transfer funds out of the bank or prime broker. Double/Triple Signature funds require two or three signatures, one of which is internal and one external. External Signature funds require only external signatures. Double Internal funds require at least two internal signatures. Single Internal funds require only one internal signature. Who Prices Portfolio identifies who values each of the assets under management. Dual/Triple Entity Pricing funds use at least one internal service and one external service to price the portfolio. External Pricing funds use only external services to price the portfolio. Collaborative Pricing funds use an external pricing service that collaborates with the manager. Internal Pricing funds price the portfolio in house. Least Objective Source of Prices provides the sources used to price individual investment positions in the portfolio. Who Sets NAV identifies whether the manager has involvement in the reporting of the NAV to fund investors. Auditor and Administrator are ranked according to Institutional Investor's Alpha ranking of hedge fund service providers.

The index values in Panel B are based on the values from Panel A. Internal Control Index is the sum of the six presented indices. The mean of the Internal Control Index is not equal to the sum of the means of its components due to missing values.

There is also substantial variation as to who prices the funds' portfolios. We observe that 9.2 percent of funds use at least one internal service and one external service to price investment positions, whereas 62.9 percent use only external services to price the portfolio. We also find that 11.7 percent of funds use collaborative pricing, whereby an external pricing service collaborates with the fund manager to determine the value of the portfolio. Finally, 16.2 percent of the funds in the sample price the portfolio internally.

The HFDD reports list all price sources that the funds use to value their invested assets. Given that some funds use multiple pricing sources, we report all sources and the least objective source used by each fund. The most objective source, exchange quotes, is used (solely used) by 69.9 percent (34.8 percent) of the funds. Over-The-Counter (OTC) quotes are used (the fund's least objective source) for 15.8 percent (5.9 percent) of the funds, whereas dealer quotes are used (the fund's least objective source) for 40.2 percent (33.2 percent). Regarding proprietary based valuation sources, we find 14.8 percent (12.7 percent) of our sample uses (as their least objective source) model-based prices. Finally, 13.4 percent of our sample generates price internally. Internal pricing is likely to be advantageous if the asset if difficult to value, allowing managers to use their superior judgment and information to provide the best estimate of the investment value. However, it provides managers with greater opportunities to manipulate reported performance. We use the fund's least objective price source for our empirical analyses.

The entity that sets and reports to investors the NAV of the hedge fund's investment portfolio differs from that which prices the individual investment positions. Pricing is done on an asset-by-asset basis and occurs on a routine, frequent basis, whereas NAV represents the net asset value of the entire portfolio and is, therefore, an aggregation of the individual prices determined by the fund's pricing systems. NAV is typically estimated and reported to investors at month end. The majority of funds (85.9 percent) have no manager involvement in the determination of the NAV, which represents the most unbiased approach to its calculation. However, the remaining 14.1 percent have some manager involvement, with most giving the manager sole determination of the setting and reporting of NAV.

We now describe the use of service providers by our funds. All funds in the sample are audited, but there is heterogeneity in auditor quality and reputation. Whereas a Big 4 auditor is typically used in accounting research to represent auditor quality or reputation (Hogan 1997; Mansi et al. 2004; Fortin and Pittman 2007), we use hedge fund industry-specific rankings of audit firms obtained from Institutional Investor's Alpha Survey, because hedge fund auditing requires specialized skills that differ from those required to audit publicly traded firms.[23] The Alpha survey is based on voting by industry participants, and is similar to Institutional Investor's ranking of equity analysts used by researchers (Stickel 1992; Hong and Kubik 2003). Of the funds in the sample, 77.0 percent are audited by an auditor that is ranked in the Alpha survey. We also use the Alpha survey to determine the quality of the fund administrator and find that 31.6 percent of sample use a ranked administrator.[24]

Table 2, Panel A also reports internal control use for all domiciles with more than ten observations. We observe variation across all the internal controls investigated within domicile group, suggesting that funds have discretion on all these internal control choices. Further, examining the regulation of our sample domiciles, we find no mandatory requirements that would require internal control adoption investigated in this study (ABA 2005; PricewaterhouseCoopers 2006). For example, the Cayman Islands have a requirement that hedge fund service providers are from an

[23] The correspondence between Big 4 and the hedge fund industry-specific measure is 85 percent.

[24] We also considered a prime broker as a potential service provider. We found, however, that almost all the funds in our sample utilized a prime broker that was ranked by the Alpha survey.

accredited registered list, but this list includes many service providers not ranked by Alpha Investors. Consistent with this fact, we observe non-ranked auditors and administrators used by Cayman Islands domiciled funds. Further, we observe a higher likelihood of engaging a ranked auditor and administrator across non-Cayman offshore funds than in U.S. domiciled funds, suggesting that any offshore findings are not solely attributable to being domiciled in the Cayman Islands.

Within each group, we numerically rank each internal control by quality. These ranks are provided in Table 2, Panel A. Auditor and administrator ranks are based on the Alpha Magazine rankings. For the remaining controls, we use HFDD's rank ordering, which is based on its determination of each control's objectivity. HFDD includes these rank orderings in the due diligence reports it prepares for its clients. Furthermore, to measure the overall association between more effective internal controls and their determinants, for each fund we sum the ranks of the individual internal controls to form an overall index score.[25] An advantage of this approach is that it provides a measure of overall internal control use while allowing for substitutes and complements. Table 2, Panel B presents descriptive statistics for the summary measures for each group and our overall internal control index.

Table 3 reports the correlations between each of the internal controls investigated, the overall internal control measure, the independent variables, and hedge fund fees. The six internal controls we examine are all significantly positively correlated (at $p < 0.10$, two-tailed), aside from the signature-administrator pairing. We also find that better internal controls are significantly positively correlated with funds and managers that are offshore, levered funds, and management and performance fees, and negatively correlated with fund age and short bias in investment strategy.

Determinants of Hedge Fund Internal Controls

Table 4 presents regressions modeling the determinants of internal controls used by the funds in the sample. To represent fund characteristics, we use the following independent variables: the natural logarithm of the fund's assets under management, the natural logarithm of the fund's age in days, and indicators for whether the fund and fund manager are offshore and for whether the fund is offshore but the manager is onshore. To proxy for portfolio characteristics and investment style, we include indicator variables for whether the fund uses leverage, whether the portfolio has a long bias, and whether the portfolio has a short bias. To control for the asset class and liquidity of the portfolio's investments, we include indicator variables for typical holding periods of portfolio investments, indicator variables for typical number of positions held by the fund, and year fixed effects.[26] Note that variables related to investment style and liquidity also control for variations in fund complexity.

Although previous research that investigates the liquidity of hedge fund investment portfolios employ a measure based on the serial correlation of self-reported monthly fund returns (Getmansky et al. 2004), we use the series of indicators representing typical number and length of positions to avoid a mechanical association between the serial correlation measure and internal controls. Namely, if internal controls affect the serial correlation of reported performance by influencing the

[25] In unreported results, we attempted to reduce the dimensionality of the internal control variables through principal component factor analysis. Although the six internal controls were positively correlated with each other, as shown in Table 3, the largest factor obtained explained only 33.9 percent of the overall variance in our internal controls.

[26] In unreported analyses we investigate whether the internal control index increases through our sample period. We find that the mean level of the index monotonically increases each year from 2002 through 2007, implying a secular improvement in hedge fund internal controls.

TABLE 3

Correlation Matrix[a]

	1	2	3	4	5	6	7	8
1. Signatures Index								
2. Who Prices Index	**0.37**							
3. Source of Prices Index	**0.09**	**0.24**						
4. NAV No Manager	**0.19**	**0.54**	**0.14**					
5. Ranked Auditor	**0.22**	**0.13**	**0.10**	**0.13**				
6. Ranked Administrator	0.07	**0.15**	**0.09**	**0.18**	**0.16**			
7. Internal control	**0.64**	**0.72**	**0.68**	**0.50**	**0.35**	**0.34**		
8. Ln(Fund Age)	−0.02	−0.04	**−0.13**	−0.07	0.05	−0.04	**−0.12**	
9. Off. Mgr. and Fund	**0.17**	**0.19**	0.08	**0.21**	**0.15**	**0.21**	**0.25**	0.01
10. Fund offshore	0.03	−0.01	−0.04	0.05	0.07	−0.03	−0.01	−0.07
11. Leverage	**0.12**	**0.13**	**0.12**	0.03	**0.13**	**0.16**	**0.17**	0.00
12. Short bias	0.02	−0.05	−0.06	−0.07	−0.03	**−0.11**	**−0.09**	0.00
13. Long bias	0.02	0.01	**−0.11**	0.06	−0.02	**0.16**	−0.02	0.05
14. Ln(Fund AUM)	**0.12**	0.04	**−0.17**	0.02	**0.17**	**0.16**	0.02	**0.49**
15. Management fee	0.05	**0.15**	**0.12**	**0.13**	**0.09**	0.04	**0.16**	−0.05
16. Performance fee	0.07	0.06	**0.36**	0.04	0.08	0.07	**0.23**	0.02
17. Restatement	−0.02	−0.02	0.05	0.00	0.04	−0.06	−0.01	0.04

	9	10	11	12	13	14	15	16
10. Fund offshore	**−0.71**							
11. Leverage	0.08	−0.02						
12. Short bias	−0.04	0.03	0.01					
13. Long bias	0.05	0.00	**0.18**	**−0.22**				
14. Ln(Fund AUM)	**0.09**	0.04	**0.11**	−0.03	**0.17**			
15. Management fee	**0.17**	−0.04	**0.09**	−0.04	0.06	−0.02		
16. Performance fee	0.01	**0.10**	**0.13**	0.01	−0.04	0.05	**0.09**	
17. Restatement	−0.01	0.06	−0.04	−0.01	−0.06	0.01	**−0.15**	0.00

[a] Variables are defined in Tables 1 and 2. Correlations in bold are statistically significant at the p=0.10 level.

ability of the manager to smooth reported performance, then the serial correlation measure of liquidity is likely to be correlated with the quality of the fund's internal controls.[27]

Table 4, Panel A presents probit regressions for the signatures required to transfer funds from the bank accounts and prime broker accounts.[28] We find that single internal signature authority is significantly less likely to be used by offshore funds, regardless of manager domicile, consistent with offshore funds providing higher quality internal controls. Larger funds are also more likely to

[27] The strength of internal controls can have an analogous effect on observed return volatility and observed tracking error. These effects imply that observed return volatility and tracking error is likely to be an outcome, in addition to a determinant, of the fund's internal controls. Given the potential reverse causality between these variables and internal controls, we do not include observed return volatility or observed tracking error as determinants of internal controls.

[28] Unlike an ordered probit, the probit regressions are not constrained to ordering assumptions related to internal control quality. An alternative approach is to present a multinomial probit for each pair-wise combination of internal controls. One issue with the multinomial specification is that the reported findings (coefficients) are presented as a difference between each alternative internal control, resulting in many reported results. For example, for pricing sources (Table 4, Panel C), there would be ten potential differences (columns).

TABLE 4

Determinants of Internal Controls

Panel A: Determinants of Signatures Required to Transfer Funds Out of Bank or Prime Broker[a,d]

	Single Internal (Probit)		Double Internal (Probit)		Single External (Probit)		Dual/Triple (Probit)		Signatures Index (Ordered Probit)	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Ln(Fund Age)	0.059	0.067	0.021	0.068	−0.010	0.081	−0.066	0.064	−0.061	0.051
Off. Mgr. and Fund	−1.122***	0.239	−0.239	0.241	1.500***	0.277	0.163	0.240	0.776***	0.200
Fund offshore only	−0.600***	0.224	−0.259	0.223	0.711***	0.268	0.597***	0.228	0.721***	0.205
Leverage	−0.215	0.156	0.003	0.159	0.080	0.172	0.143	0.159	0.173	0.125
Short bias	0.219	0.196	0.024	0.196	−0.475**	0.207	0.154	0.182	−0.036	0.163
Long bias	0.089	0.172	0.157	0.164	−0.405**	0.204	0.103	0.162	−0.042	0.130
Ln(Fund AUM)	−0.119**	0.051	0.030	0.060	0.049	0.054	0.056	0.053	0.084*	0.046
Years	0.028	0.246	0.178	0.242	0.598**	0.278	−0.588**	0.231	−0.294	0.212
Quarters	0.159	0.244	0.132	0.232	0.441	0.276	−0.552**	0.236	−0.348	0.214
Weeks	−0.334	0.329	−0.351	0.330	0.843**	0.373	−0.100	0.303	0.178	0.266
Days	−0.121	0.317	0.077	0.295	0.547	0.340	−0.409	0.285	−0.116	0.239
1000+ positions	−0.001	0.497	1.249***	0.452	−1.407**	0.586	−0.632	0.475	−0.606**	0.285
200–999 positions	−0.142	0.295	−0.022	0.271	0.033	0.299	0.046	0.252	0.072	0.197
100–199 positions	0.077	0.246	0.436*	0.236	−0.282	0.242	−0.208	0.244	−0.229	0.186
1–39 positions	0.081	0.175	0.113	0.179	−0.157	0.178	0.023	0.160	−0.054	0.135
Intercept	1.741**	0.782	−1.020	0.957	−3.029***	0.977	−1.598*	0.820		
Year fixed effects	Yes		Yes		Yes		Yes		Yes	
Pseudo R^2	0.128		0.088		0.210		0.071		0.051	
Log-Likelihood	−192.937***		−200.654**		−179.301***		−220.259***		−529.657***	
n	403		403		403		403		403	

(continued on next page)

Panel B: Determinants of Who Prices the Portfolio[b,d]

Independent Variables	Manager Only (Probit)		Manager and Administrator (Probit)		Administrator Only (Probit)		Dual/Triple Entity (Probit)		Who Prices Index (Ordered Probit)	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Ln(Fund Age)	0.105	0.082	0.034	0.080	−0.071	0.065	0.031	0.095	−0.054	0.051
Off. Mgr. and Fund	−1.261***	0.264	−0.003	0.278	0.861***	0.225	−0.072	0.404	0.678***	0.211
Fund offshore only	−0.706***	0.238	0.132	0.267	0.461**	0.214	0.047	0.367	0.508**	0.207
Leverage	−0.305*	0.170	−0.027	0.177	−0.069	0.151	0.698***	0.223	0.346***	0.126
Short bias	0.493**	0.208	0.168	0.198	−0.550***	0.178	0.422	0.271	−0.168	0.175
Long bias	0.046	0.182	0.035	0.188	0.059	0.168	−0.260	0.233	−0.101	0.125
Ln(Fund AUM)	0.003	0.063	−0.022	0.076	−0.038	0.062	0.120	0.094	0.039	0.048
Years	0.060	0.261	−0.296	0.273	0.150	0.213	0.054	0.337	0.090	0.203
Quarters	0.168	0.267	−0.113	0.243	−0.082	0.197	0.171	0.350	0.035	0.213
Weeks	−0.480	0.384	−0.506	0.409	0.342	0.295	0.372	0.467	0.552*	0.317
Days	−0.196	0.372	−0.212	0.353	0.084	0.278	0.250	0.429	0.199	0.261
1000+ positions	0.817	0.539	0.861**	0.437	−0.713	0.441	NA	NA	−0.995***	0.343
200–999 positions	0.088	0.309	−0.167	0.345	0.281	0.259	−0.769**	0.351	−0.237	0.222
100–199 positions	−0.064	0.305	0.199	0.254	0.017	0.220	−0.205	0.318	−0.114	0.190
1–39 positions	0.323	0.204	0.047	0.206	−0.313*	0.160	0.211	0.220	−0.111	0.141
Intercept	−0.785	1.062	−0.860	1.249	0.896	1.048	−4.481***	1.585		
Year fixed effects	Yes		Yes		Yes		Yes		Yes	
Pseudo R^2	0.176		0.052		0.111		0.202		0.064	
Log-Likelihood	−147.160***		−139.920		−237.264***		−95.981***		−398.853***	
n	406		406		406		406		406	

(continued on next page)

Panel C: Determinants of Sources for Pricing the Portfolio[c,d]

	Manager (Probit)		Model (Probit)		Dealer (Probit)		OTC (Probit)		Exchange (Probit)		Source of Prices Index (Ordered Probit)	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Ln(Fund Age)	0.099	0.099	−0.081	0.074	0.086	0.066	0.010	0.072	−0.098	0.070	−0.052	0.055
Off. Mgr. and Fund	−0.764***	0.289	−0.571**	0.283	0.376	0.237	0.178	0.293	0.501**	0.226	0.366*	0.201
Fund offshore only	−0.533*	0.293	−0.587**	0.256	0.704***	0.223	0.331	0.287	0.339	0.213	0.087	0.189
Leverage	−0.483**	0.209	−0.294	0.180	0.393***	0.152	0.254	0.186	−0.189	0.161	0.121	0.127
Short bias	0.472**	0.239	0.141	0.190	0.489***	0.189	−0.204	0.226	−0.714***	0.203	−0.374**	0.152
Long bias	0.485**	0.229	−0.218	0.199	0.557***	0.158	0.049	0.177	0.113	0.174	−0.317**	0.133
Ln(Fund AUM)	0.182**	0.083	0.139*	0.073	0.078	0.075	0.072	0.084	0.006	0.055	−0.119*	0.071
Years	1.534***	0.382	0.227	0.240	0.210	0.234	−0.672**	0.297	−1.086***	0.259	−0.857***	0.175
Quarters	0.636	0.405	−0.098	0.267	0.253	0.234	−0.035	0.276	−0.535*	0.273	−0.221	0.171
Weeks	0.086	0.516	−0.335	0.362	0.276	0.362	0.221	0.359	0.123	0.371	−0.188	0.219
Days	0.521	0.543	−0.615	0.431	0.121	0.303	−0.118	0.369	0.600	0.392	0.196	0.242
1000+ positions	1.211**	0.528	1.225**	0.512	0.226	0.368	0.123	0.502	0.085	0.501	−0.730**	0.372
200−999 positions	−0.272	0.572	0.380	0.310	−0.278	0.271	−0.300	0.283	−0.134	0.298	0.249	0.206
100−199 positions	0.459	0.358	0.161	0.267	−0.224	0.214	−0.407	0.260	0.624**	0.297	0.103	0.188
1−39 positions	0.716***	0.235	0.103	0.226	−0.308*	0.160	−0.146	0.184	−0.652***	0.185	−0.237*	0.138
Intercept	−7.482***	1.734	−2.834**	1.152	−2.776**	1.275	−2.234	1.404	1.585*	0.874		
Year fixed effects	Yes		Yes		Yes		Yes		Yes		Yes	
Pseudo R^2	0.359		0.126		0.166		0.152		0.232		0.088	
Log-Likelihood	−103.054***		−142.302***		−226.988***		−149.837***		−184.140***		−514.151***	
n	405		405		405		405		405		405	

(continued on next page)

Panel D: Determinants of NAV Independence and Service Providers (Probit)[d]

Independent Variables	NAV No Manager		Ranked Auditor		Ranked Administrator	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Ln(Age Fund)	−0.034	0.079	0.016	0.068	−0.137**	0.069
Off. Mgr. and Fund	1.452***	0.282	1.021***	0.237	1.270***	0.298
Fund Offshore Only	0.958***	0.246	0.845***	0.226	0.924***	0.299
Leverage	−0.054	0.183	0.376***	0.162	0.234	0.161
Short Bias	−0.295	0.221	−0.137	0.188	−0.447**	0.199
Long Bias	0.142	0.197	−0.219	0.172	0.347**	0.163
Ln(Fund AUM)	−0.037	0.069	0.109*	0.057	0.123**	0.059
Years	−0.109	0.251	−0.130	0.234	−0.017	0.232
Quarters	0.235	0.267	0.120	0.251	−0.089	0.229
Weeks	0.033	0.359	−0.165	0.304	0.244	0.288
Days	−0.132	0.334	−0.313	0.296	−0.077	0.305
1000+ Positions	−0.100	0.569	−0.665	0.446	0.106	0.426
200–999 Positions	−0.222	0.272	−0.102	0.301	−0.288	0.276
100–199 Positions	−0.062	0.299	−0.063	0.267	0.228	0.222
1–39 Positions	0.002	0.212	−0.132	0.175	−0.242	0.165
Intercept	0.855	1.148	−1.959**	0.913	−2.980***	0.912
Year Fixed Effects	Yes		Yes		Yes	
Pseudo R^2	0.171		0.131		0.143	
Log-Likelihood	−135.054***		−191.070***		−219.868***	
n	406		407		407	

(continued on next page)

Panel E: Determinants of Overall Internal Controls[d] (Ordinary Least Squares)

Independent Variables	Coefficient	Std. Error
Ln(Fund Age)	−0.229*	0.119
Off. Mgr. and Fund	2.727***	0.469
Fund Offshore Only	1.973***	0.468
Leverage	0.593**	0.283
Short Bias	−0.992***	0.379
Long Bias	−0.406	0.306
Ln(Fund AUM)	0.036	0.115
Years	−1.460***	0.450
Quarters	−0.606	0.446
Weeks	0.422	0.600
Days	−0.043	0.528
1000+ Positions	−2.259***	0.861
200–999 Positions	−0.053	0.441
100–199 Positions	−0.189	0.407
1–39 Positions	−0.547*	0.303
Intercept	6.817***	1.903
Year Fixed Effects	Yes	
Adjusted R^2	0.224	
F-stat	5.792***	
n	403	

*, **, *** Denotes significance at p = 0.10, p = 0.05, and p = 0.01, respectively (based on a two-tailed test).

[a] Individual intercepts not reported on ordered probit. Independent variables defined in Table 1. Dependent variables defined in Table 2. Standard errors for all regressions are clustered at the manager level.

[b] Individual intercepts not reported on ordered probit. Independent variables defined in Table 1. Dependent variables defined in Table 2. Standard errors for all regressions are clustered at the manager level. NA represents missing variable due to the model being perfectly identified with its inclusion.

[c] Individual intercepts not reported on ordered probit. Independent variables defined in Table 1. Standard errors for all regressions are clustered at the manager level. Dependent variables defined in Table 2.

[d] Independent variables are defined in Table 1. Omitted indicator variables are Months and 40–99 Positions. Standard errors for all regressions are clustered at the manager level. Dependent variables are defined in Table 2.

use single signature authority. As shown from the External Only and Dual/Triple Entity models, offshore funds and offshore managers are also more likely to require external signatures. Transaction costs appear to affect signature authority, with funds that hold more than 1,000 positions less likely to require external signatures. If these funds move positions regularly and are relatively liquid, then the requirement to have an external party authorize the transfer of funds on each occasion could be overly burdensome and costly. These funds are, however, more likely to require at least two internal signatures, because large numbers of positions increase the likelihood of fraud.

The right-most column of Table 4, Panel A presents an ordered probit regression in which the dependent variable takes on the index values (0, 1, 2, 3), representing Single Internal, Double Internal, External Only, and Dual/Triple Entity signature authority, respectively, enabling us to determine the overall association between more effective signature internal controls and the independent variables. Under this specification we observe that larger funds and funds that are domiciled offshore are significantly associated with increased control over signatures required to transfer funds out of bank and prime broker accounts.

Table 4, Panel B presents a series of probit regressions, modeling who prices the hedge fund's investment positions. We find that offshore funds are significantly less likely to internally price their portfolios and significantly more likely to use external pricing services. Levered funds are also significantly less likely to have the manager price investment positions and more likely to use external pricing services. Furthermore, funds with a short bias in their investment positions are more likely to use internal pricing and less likely to use external pricing. This finding is likely driven by the proprietary nature of short selling and the potential for outsiders to profit from knowing that the fund has to unwind its short positions. The right-most column of the panel presents an ordered probit regression of overall quality with respect to who prices individual investment positions. The coefficients in this specification are consistent with the inferences from the individual pricing regressions, with offshore funds and levered funds more likely to rely on external rather than internal pricing, and those funds with a short bias in their investment strategy more likely to rely on internal pricing than external pricing.

Table 4, Panel C presents probit regressions modeling the source of prices used to value the funds' portfolios. Consistent with the results for who prices the portfolio, offshore funds are less likely to use internal and model-based sources of prices and more likely to use more dealer and exchange-based sources to value their portfolios. Funds that use leverage are less likely to use internal sources and more likely to use dealer sourced prices, consistent with hedge funds with greater potential agency costs using relatively more objective pricing sources. Further, we observe that funds with some bias in their investment style, long or short, are less likely to use external sources, such as dealer and exchange-based prices, and are more likely to use internal price sources. The right-most column of Panel C provides the ordered probit regression of overall pricing source quality. We again observe that funds with bias in their investment style making less use of objective price sources, and some evidence that small funds use less objective sources. We also observe that funds with managers domiciled offshore make greater use of more objective pricing source measures.

The left-most column of Table 4, Panel D presents probit regressions modeling whether the NAV is determined without manager involvement. We observe that managers of offshore funds, regardless manager location, are more likely to determine and report NAV without manager involvement. These findings are consistent with the pricing findings observed in Panels B and C. The remaining columns from Panel D model whether the fund uses an auditor or an administrator ranked by Institutional Investor's Alpha Survey. Larger funds and offshore funds are more likely

to use a higher quality auditor and administrator. Additionally, funds that use leverage are more likely to use a higher quality auditor, and younger funds are more likely to use a higher quality administrator.

Table 4, Panel E provides the ordinary least-squares regression modeling overall internal control quality as a function of fund characteristics. Overall, the evidence is consistent with offshore funds and managers adopting stronger internal control mechanisms to decrease the likelihood of fraud and financial misstatements, incorporating stricter signature authority for the transfer of funds, using independent pricing mechanisms, and employing more reputable outside service providers. These findings suggest that investors demand greater internal controls for funds that are not subject to the U.S. legal infrastructure. There is also evidence that funds that use leverage, thereby increasing fund risk, also employ stronger internal control mechanisms. Overall, internal control quality is negatively associated with fund age, consistent with younger funds having greater incentives to implement stronger internal controls to reduce greater concerns with adverse selection and moral hazard. Finally, short bias funds have lower quality internal controls, driven in part by these funds reducing proprietary costs that can arise from external parties learning the funds' investment positions.

Determinants of Hedge Fund Fees

Managers have incentives to deflate, inflate, and smooth reported performance and assets under management. First, when the investor cashes out, the value of the investor's asset is revealed by the hedge fund. At this point, the manager has an incentive to deflate the revealed value of the investor's claim, thereby increasing the value of assets retained by the hedge fund. Managers also have incentives to inflate reported performance because they collect management fees based on reported assets under management and reported performance is a determinant of capital flows (Fung et al. 2008). Finally, they have incentives to smooth reported performance to reduce the perceived volatility of the fund and therefore increase its perceived risk adjusted performance (Getmansky et al. 2004).

We posit that more extensive internal controls improve the accuracy and precision of asset valuations and returns disclosures made to investors, and investors are willing to pay higher management fees to funds with more extensive internal controls because of the decreased likelihood and magnitude of losses arising from managers' manipulation of reported performance. Therefore, *ceteris paribus*, we predict that internal controls that provide greater opportunity to manipulate reported performance are negatively associated with the management fees paid to fund managers.

Aside from any variations in management fees, the use of internal controls can also affect performance fees. The incentives to manipulate the reported performance and valuations increase with the sensitivity of the manager's compensation relative to reported values and performance. An obvious mechanism to vary this sensitivity is through the proportion of fees rewarded from investment profits, namely the performance fee. Therefore, *ceteris paribus*, we predict that internal controls that provide greater opportunity to manipulate reported performance will be negatively associated with the performance fees paid to fund managers.

Table 5 presents regressions of the determinants of hedge fund fees. For these regressions, we use the same independent variables as in the "Determinants of Hedge Fund Internal Controls" section.[29] Panel A presents ordinary least-squares regressions that model the determinants of the

[29] The use of the same independent variables in the determinants of internal control use and the hedge fund fee models is equivalent to using a two-stage approach, in which the first stage estimates internal control use, and the second stage regresses fees on the residuals from the first stage.

TABLE 5

Determinants of Hedge Fund Fees

Panel A: Determinants of Management Fees (Ordinary Least Squares)

	Controls Only		With Index		With Restatement	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Internal Control Index			0.005	0.011		
Restatement					−0.227**	0.091
Ln(Fund Age)	−0.013	0.021	−0.013	0.021	−0.008	0.022
Off. Mgr. and Fund	0.208***	0.075	0.188**	0.080	0.223***	0.075
Fund Offshore Only	0.116	0.078	0.095	0.081	0.132*	0.078
Leverage	0.074	0.049	0.069	0.050	0.066	0.048
Short Bias	−0.053	0.058	−0.046	0.060	−0.061	0.056
Long Bias	0.019	0.058	0.021	0.060	0.016	0.058
Ln(Fund AUM)	−0.002	0.015	−0.002	0.015	−0.004	0.015
Years	−0.087	0.063	−0.082	0.063	−0.070	0.062
Quarters	0.039	0.061	0.040	0.061	0.047	0.060
Weeks	0.205**	0.104	0.204*	0.104	0.204*	0.104
Days	0.088	0.099	0.083	0.100	0.087	0.098
1000+ Positions	−0.398*	0.240	−0.378	0.236	−0.393	0.244
200–999 Positions	0.037	0.091	0.043	0.091	0.037	0.093
100–199 Positions	−0.121*	0.065	−0.114*	0.066	−0.096	0.066
1–39 Positions	0.021	0.054	0.025	0.056	0.015	0.054
Intercept	1.387***	0.231	1.364***	0.250	1.408***	0.240
Year Fixed Effects	Yes		Yes		Yes	
Adjusted R^2	0.070		0.066		0.086	
F-stat	3.275***		3.062***		3.934***	
n	406		402		404	

(continued on next page)

Panel B: Determinants of Performance Fees (Ordinary Least Squares)

	Controls Only		With Index		With Restatement	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Internal Control Index			0.261**	0.102		
Restatement					−0.137	0.591
Ln(Fund Age)	0.135	0.236	0.193	0.236	0.142	0.243
Off. Mgr. and Fund	1.303	0.887	0.516	0.782	1.319	0.890
Fund Offshore Only	1.640*	0.849	1.008	0.776	1.649*	0.852
Leverage	0.452	0.416	0.273	0.420	0.435	0.415
Short Bias	0.013	0.427	0.347	0.447	−0.010	0.432
Long Bias	−0.632	0.522	−0.473	0.528	−0.625	0.527
Ln(Fund AUM)	−0.014	0.112	−0.028	0.114	−0.017	0.113
Years	−2.698***	0.609	−2.276***	0.581	−2.688***	0.614
Quarters	−0.880*	0.450	−0.715	0.453	−0.890*	0.454
Weeks	−0.095	0.485	−0.201	0.501	−0.091	0.487
Days	1.525*	0.892	1.557*	0.901	1.518*	0.903
1000+ Positions	−0.279	2.455	0.324	2.426	−0.275	2.459
200–999 Positions	−0.251	0.614	−0.230	0.631	−0.248	0.614
100–199 Positions	−0.766*	0.443	−0.688	0.437	−0.747*	0.448
1–39 Positions	−1.950***	0.522	−1.782***	0.498	−1.967***	0.525
Intercept	19.633***	1.967	18.064***	2.136	19.672***	2.073
Year Fixed Effects	Yes		Yes		Yes	
Adjusted R^2	0.167		0.181		0.165	
F-stat	2.288***		2.144***		2.170***	
n	406		402		404	

*, **, *** Denotes significance at p = 0.10, p = 0.05, and p = 0.01, respectively, (based on a two-tailed test).

Variables are defined in Tables 1 and 2. Omitted indicator variables are Months and 40–99 Positions. Standard errors for all regressions are clustered at the manager level.

management fee paid to the hedge fund manager, with the left-most model presenting the results with controls only, and the second model with controls and our measure of internal control quality. For ease of interpretation, we multiply the dependent variable, the management fee, by 100. Examining our overall measure of internal controls on management fees reveals that there is no significant effect. Taken by itself, this evidence suggests that our investigated internal controls are not priced through the management fees.

Table 5, Panel B presents ordinary least-squares regressions that model the determinants of the performance fee paid to the manager. As with the management fee, we multiply the performance fee by 100. With respect to the control variables, performance fees are positively associated with funds that have higher turnover in their investment portfolios. Examining the internal control coefficients, we find that better internal controls are significantly positively associated with performance fees ($\beta = 0.261$; $p < 0.05$). In terms of economic significance, the standard deviation of the index is 2.78, which approximately represents going from internal pricing to dual/triple entity pricing. In terms of fees, a one standard deviation increase in the index is associated with a 0.726 percentage point increase in performance fee.

The significant coefficient on the internal control index is consistent with investors protecting against potential misstatement as a result of inadequate internal controls through lower performance fees. It is also consistent with investors and hedge funds reducing the costs from moral hazard and the motivation for managers to manipulate reported performance.[30] This finding suggests that, as the opportunity for manipulation of fund performance increases, the reliance on fund performance as an incentive mechanism decreases.

Endogeneity

A potential concern is that our internal control measures are endogenous. Although we include the variables that predict internal control quality in the model determining fees, we recognize that endogeneity could affect our inferences because of simultaneity and/or correlated omitted variables. Below we discuss these endogeneity concerns and how we address them.

First, our motivation for adopting our empirical design is driven by the ordering of the various parameters investigated in this study. Given that fund managers have specialized expertise in particular areas of investment (Brown and Goetzmann 2003), it is unlikely that they change fund characteristics over time, justifying the assumption that these characteristics are predetermined.[31] Internal controls are then likely chosen as a function of these fund characteristics. The management and performance fees are detailed in the PPM to investors. Investors first observe these predetermined characteristics of the fund and the internal controls used by the fund, and then decide whether to invest in the fund based on the terms and fees offered. Furthermore, given that the observed internal controls and fees are determined based on those offered to all fund investors, the likelihood that internal controls and fees vary by individual investor characteristics is low.[32]

---

[30] We also perform a bootstrap analysis using an equal sample of funds with 20 percent performance fees randomly selected without replacement with the non-20 percent performance fee funds ($n = 67$) to examine the extent to which our findings and power could be influenced by the high proportion of funds with 20 percent performance fees. This selection process was repeated for 1,000 iterations to estimate the parameter and level of significance. We again observe a significantly positive association between internal controls and performance fees, suggesting our results are invariant to this alternative sampling procedure.

[31] Moreover, investors have the ability and the incentives to monitor the fund's internal controls on an ongoing basis, thereby limiting a manager's ability to weaken the fund's internal controls.

[32] We acknowledge that some investors have greater bargaining power and are, therefore, able to negotiate special fees and terms. But, such fees and terms are laid out in "side letters" that are separate from the PPM. Further, some investors could negotiate a "favored nation" clause, which would result in any favorable terms contracted with other investors also granted to them, suggesting that the presence of such a clause would increase the overall cost to the fund when offering more favorable terms to some investors.

Second, to investigate the likelihood that unobserved fund characteristics determine both internal controls and fees, we adopt the approach developed by Rosenbaum (2002) to estimate a bound on the extent to which omitted variables would have to be correlated with our measures to drive the results. We use Krauth's (2007) calculation to estimate how large the correlation between unobserved heterogeneity and the internal control index would have to be compared to the correlation between the internal control index and the other independent variables. For the 95 percent confidence interval on the internal control index coefficient to include zero, the correlation between unobserved heterogeneity and the index would have to be greater than 56 percent of the correlation between the index and the other independent variables. Although such a correlation is possible, we believe it to be unlikely given the range of fund characteristics included in our independent variables and the inclusion of relevant variables used in published hedge fund research. Nevertheless, the extent to which such omitted variables exist is a limitation of our study.

Third, an alternative approach to address potential endogeneity when examining the determinants of internal controls and fees is to hold the investment characteristics of the fund constant, thereby ruling out that the internal controls are determined by, or are capturing, the underlying investment characteristics of the fund. To implement this approach, we perform our analyses only for those funds that typically held 1 to 39 positions, the investment positions category with the greatest number of sample funds. For these 172 funds, we observe the determinants of internal controls and the association between internal controls and performance fees are consistent with the presented results for the full sample ($\beta = 0.502$; se $= 0.147$; $p < 0.01$). Additionally, we perform our analyses only for those hedge funds that typically held their positions for one year and greater, the holding period category with the greatest number of sample funds. Again, we observe the determinants of internal control use and the association between internal controls and fees to be consistent with the overall sample ($\beta = 0.622$; se $= 0.231$; $p < 0.01$).

Finally, we consider an alternative measure based on outcomes, namely restatements. Restatements are not chosen by hedge funds at the time of due diligence, making them predetermined. Examining the right-most column of Table 5, Panel A, we observe a statistically significant negative relationship between management fees and whether the fund has restated the performance reported to investors ($\beta = -0.227$; $p < 0.05$). This relationship between restatements and management fees is also economically significant. For the mean fund, a restatement is associated with $692,000 less in management fees per annum, which is equivalent to a 15 percent ($692,000/$4,608,633) reduction. The observation of restated performance by the hedge fund increases the likelihood that future fund performance be restated, thereby reducing confidence in the accuracy of present and future reported fund performance. Therefore, this evidence demonstrates a significant association between restatements, a predetermined and observable event that occurred prior to contracting, and management fees paid by investors, providing further support for the conclusion that investors protect against the risk of future misstatements by paying lower fees. However, in contrast to the management fee findings, we find little evidence of a relationship between restatements and performance fees.

VI. ANALYSIS OF FRAUD INVESTIGATIONS

To investigate a performance consequence from the use of internal controls, we examine whether internal controls are associated with future regulatory investigations of fraud and/or financial misstatements on the part of the fund and its managers.[33] We use investigations instead

[33] We do not investigate the association between internal controls and risk-adjusted returns because their predicted association is ambiguous. While stronger internal controls could lead to better performance for hedge fund investors, prior research finds evidence consistent with hedge fund managers intentionally smoothing self-reported returns to increase observed risk-adjusted performance (for a discussion, see Getmansky et al. 2004). Internal controls can reduce oppor-

of the resolution of legal proceedings because of the long duration between the commencement of an investigation and the conclusion of legal proceedings. Such investigations can be considered an extreme form of poor performance because they are typically associated with fund liquidations and deeply discounted investor redemptions.

For the funds in the sample we search both Factiva and the SEC's website for investigations of fraud and/or financial misstatements. For Factiva we use a keyword search that captures fraud and/or financial misstatement events and carry out the search over the period starting the month after completion of the due diligence report and ending in March 2009.[34] For the SEC we search the SEC's website for the name of every fund in our sample. We then read every Factiva article and SEC document to determine whether the fund is under investigation for fraud and/or financial misstatements. We identify 12 funds in our sample as investigated or under investigation of fraud and/or financial misstatements. These events are not trivial. All 12 funds were investigated for fraud and/or intentional financial misstatement, and all were liquidated with investors incurring substantial losses.

Table 6 presents probit regressions where the dependent variable is coded as 1 if the fund is subject to investigations, and 0 otherwise. The independent variables are the same variables we use to examine the determinants of the internal controls, thereby allowing us to investigate whether deviations from expected internal controls are associated with future allegations and/or investigations.

The first column presents the baseline probit regression model, which is statistically significant at $p < 0.001$. Levered funds are significantly less likely to be subject to future investigations (marginal effect = -1.00 percentage points), suggesting that leverage providers either carry out explicit monitoring or implement extensive screening prior to providing credit. Funds that hold their investment positions for quarters (marginal effect = -0.43 percentage points) are significantly less likely to have been investigated or under investigation for fraud and/or financial misstatements. Consistent with greater opportunities to commit fraud and/or financial misstatements, we find that funds holding thousands of investment positions are significantly more likely to be subject to future regulatory investigations (marginal effect = 5.94 percentage points). Further, the coefficient on fund age is negative and significant. Recall from Table 4, Panel E that we find that the internal control index decreases in fund age, suggesting that older funds have valuable reputations that substitute for explicit internal controls by providing incentives against carrying out actions that would lead to investigations of fraud and/or financial misstatements.

In the second column of Table 6 we introduce the internal control index. Consistent with higher levels of internal controls being associated with a lower probability of investigation, the coefficient on the index is negative, though it is not statistically significant. Next we examine the effect of individual internal controls on future regulatory investigations of fraud and/or financial misstatement. When we introduce the indicator variable for whether the manager is not involved in setting and reporting the fund's official NAV to investors, its coefficient is negative and significant ($\beta = -1.076$; $p < 0.05$), with a marginal effect of 1.14 percentage points. With respect to economic significance, the unconditional probability of a sample fund being investigated is 3 percent, implying that not having the fund manager involved in setting and reporting the NAV reduces the likelihood of being investigated for fraud by over one-third.

---

tunistic manipulation, suggesting that funds with weaker internal controls are likely to have higher (artificial) observed risk-adjusted performance. Given these conflicting predictions, we therefore examine fraud and misstatements for which there is an unambiguous prediction.

[34] We used the following keywords: wind w/5 down; fraud*; restate*; bankrupt*; death; liquidat*; blow* w/5 up; failure; fine*; defraud*; guilty; conspir*; scandal; deceive*; plea; missing; insolven*; collapse*; probe*; prosecut*; fled; and indict*.

TABLE 6

Determinants of Fraud Investigations

	Base		With Index		With NAV No Manager	
Independent Variables	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
Internal Control Index			−0.040	0.049		
NAV No Manager					−1.076**	0.403
Ln(Fund Age)	−0.252**	0.113	−0.267**	0.111	−0.309***	0.118
Off. Mgr. and Fund	0.060	0.422	0.214	0.491	0.540	0.543
Fund Offshore Only	0.298	0.426	0.471	0.499	0.766	0.535
Leverage	−1.011***	0.338	−1.024***	0.361	−1.157***	0.399
Short Bias	−0.357	0.461	−0.386	0.495	−0.314	0.463
Long Bias	0.515	0.330	0.548	0.356	0.535	0.355
Ln(Fund AUM)	−0.051	0.070	−0.049	0.070	−0.060	0.075
Years	−1.009*	0.601	−1.037*	0.580	−1.471**	0.591
Quarters	−0.856**	0.352	−0.913**	0.376	−1.001**	0.399
Weeks	−0.001	0.441	−0.001	0.444	−0.291	0.454
Days	−0.167	0.427	−0.148	0.413	−0.172	0.437
1000+ Positions	1.376**	0.599	1.228**	0.584	1.517**	0.628
200−999 Positions	0.396	0.501	0.377	0.506	0.527	0.490
100−199 Positions	−0.001	0.526	−0.076	0.518	0.104	0.547
1−39 Positions	−0.589	0.383	−0.696*	0.397	−0.738*	0.407
Intercept	−2.875**	1.111	−2.677**	1.189	−1.617	1.209
Year Fixed Effects	Yes		Yes		Yes	
Pseudo R^2	0.309		0.316		0.353	
Log-Likelihood	−32.322***		−31.952***		−30.244***	
n	401		397		400	

*, **, *** Denotes significance at p = 0.10, p = 0.05, and p = 0.01, respectively, (based on a two-tailed test).

Table reports probit estimation. Variables are defined Tables 1 and 2. Omitted indicator variables are Months and 40–99 Positions. Standard errors for all regressions are clustered at the manager level. Dependent variable is an indicator for whether there allegations and/or investigations of fraud and/or financial misstatements for the fund.

Overall, we observe some evidence consistent with the choice of internal controls being associated with a critical performance consequence, namely fraud investigations against the hedge fund. Moreover, we find that reputational incentives as proxied by fund age and explicit external monitoring and/or screening by leverage providers are associated with lower likelihoods of future regulatory investigations of fraud and/or financial misstatements.

VII. ROBUSTNESS

Unit of Analysis

We perform the analyses of the determinants of internal controls and investor fees at the fund level and cluster the standard errors by manager to address any manager-related cross-correlations. Alternatively, we could perform the analyses at the manager level. To investigate the robustness of our results, we repeat all analyses using three alternative selection criteria with only one observation per manager. The criteria are: (1) the first investigated fund by manager, (2) the last investigated fund by manager, and (3) an average of the dependent and independent variables by man-

ager. For all three approaches, the results (untabulated) for both the determinants of internal controls and their relationship to fees are consistent with the presented findings.

The reported analyses employ fund characteristics associated with assets under management and fund age, which have been commonly used by researchers to capture cross-sectional variation in fund characteristics (Aragon 2007; Brown et al. 2008). The extent to which the use of fund-level characteristics, rather than manager-level characteristics, influences our findings is an empirical question. We consider two manager-based measures, namely the manager's assets under management and the natural logarithm of the number of funds that the manager manages. The Pearson correlations between the fund's assets under management and the natural logarithm of the fund's age in days with the manager's assets under management are 0.61 ($p < 0.01$) and 0.25 ($p < 0.01$), respectively. The Pearson correlations between the fund's assets under management and the natural logarithm of the fund's age in days with the natural logarithm of the number of funds that the manager manages are 0.16 ($p < 0.01$) and 0.03 ($p < 0.56$), respectively. Using the manager characteristics in the specifications presented in Table 5, we again observe a significantly positive association between overall internal control quality and performance fees ($\beta = 0.236$; $p < 0.05$) and no significant association between internal control quality and management fees ($\beta = 0.009$; se = 0.011).

Internal Control Index

To examine the extent that the choice and weighting of index items significantly affects our inferences, we examine an alternative index to capture overall hedge fund internal control quality. In particular, we adopt an index based solely on the extent that the internal controls did or did not include manager involvement. For signature authority to transfer funds, who prices the portfolio, and the pricing source of the portfolio, we code 0 for each component if the manager was involved, and 1 if the manager was not involved. For this alternative index, we score each fund between 0 and 3 (mean = 2.010; standard deviation = 0.896). Consistent with the reported hedge fund quality index results, we again observe no significant association between internal controls and management fees ($\beta = 0.012$; se = 0.036), and a significantly positive association between internal controls and performance fees ($\beta = 0.586$; se = 0.287; $p < 0.05$). In terms of economic significance, a one standard deviation increase in the alternative index is associated with a 0.526 percentage point increase in the performance fee. The magnitude of this effect is similar to the magnitude for the reported internal control index (0.526 versus 0.726).

Self-Selection in Reporting Performance

We compare our sample hedge funds that do (n = 232) and do not (n = 195) self-report performance to the Lipper TASS database, which is the primary database used in academic research on hedge funds (see Lo [2008] for a discussion of the database). Similar to our sample, Lipper TASS and other hedge fund databases do not comprise the complete universe of hedge funds, as these databases consist of information self-reported from hedge funds that choose to be included in the databases. However, these datasets are an order of magnitude greater in number of reported funds than our database, and the extent of self-selection is helpful for evaluating the implications of our findings to the broader hedge fund population.

We perform univariate tests for differences across fund age, fund and manager location, use of leverage, investment style, management and performance fees, and internal controls. Across all these variables, the only significant difference is that older sample funds have a significantly higher likelihood of reporting to Lipper TASS ($p < 0.01$), with the mean number of days being 1,252 (742) for sample funds reporting (not reporting) to Lipper TASS. An explanation for the difference across fund age is the potential for self-selection based on performance for funds that choose to report to Lipper TASS. Specifically, if older surviving funds generally have better

performance, then they have greater incentives to self-report to database vendors for promotion and marketing purposes. It should be noted that this significant difference is above any other potential self-selection bias arising from HFDD being appointed to perform due diligence on a fund. In addition, we note that investors were actively considering an investment in the funds included in our sample, implying that the observed trade-offs reflect market conditions at the point of due diligence. Given these results, we expect our findings to be representative across a broad set of funds.

Further Controlling for Investment Style

Although we control for asset class and investment style by including typical holding period, number of positions, and the use of long or short bias, we examine the extent that further controls for asset class and investment style influences our observed findings. For our 232 sample observations that report to Lipper TASS, we include indicator variables for each fund's investment style as classified by Lipper TASS. These investment styles include: convertible arbitrage, dedicated short seller, emerging markets, equity market neutral, event-driven, fixed income arbitrage, global macro, and long/short equity. In unreported results, all findings for the determinants of internal controls and the determinants of investor fees are robust to the inclusion of the Lipper TASS investment style indicator variables and consistent with our empirical predictions, with the exception of our investment position variables. Replicating the results from Table 4, Panel E with investment style variables, we observe that both 1000+ positions ($\beta = -0.511$, originally -2.259) and 1–39 positions ($\beta = 0.243$, originally -0.547) are no longer significantly negatively associated with our overall internal control index. One interpretation of these findings is that our investment position variables are, to some extent, capturing the underlying investment style of the funds, as based on Lipper TASS classifications. Finally, we note that the coefficient on 1000+ positions ($\beta = -0.985$, originally -0.606) remains significant and negatively associated with greater signature authority with the inclusion of investment style variables (Table 4, Panel A).

Offshore Results

In the reported analyses, we categorize the domicile of the hedge fund into on- and offshore. We examine the robustness of our findings by replacing the fund domicile variables in the reported analyses with indicator variables for each country with ten or more funds. Our results are not altered by the inclusion of country indicator variables.

VIII. CONCLUSION

We investigate the determinants and internal controls of hedge funds and their association with the fees that funds charge investors using proprietary due diligence reports commissioned by sophisticated investors. The multifaceted nature of this data provides substantial detail regarding fund characteristics, operations, and internal controls. These reports are based on interviews, analyses of contract terms and financial position, and third-party sources, in contrast to other hedge fund data sets that are solely based on self-reported, unverified information.

Consistent with our predictions, we find that managers of funds domiciled offshore adopt stronger mechanisms to decrease the likelihood of fraud and financial misstatements, incorporate stricter signature authority to transfer funds, use external pricing services, and use more reputable outside service providers. We also observe that levered funds and younger funds are more likely to have stronger internal controls. Further, we find that funds that have a short bias investment strategy are less likely to employ independent pricing sources. This finding likely reflects greater proprietary costs from others trading against the short positions of these funds.

With respect to outcomes from internal control weaknesses, we find that funds that have restated performance charge lower management fees. Further, we find a positive association be-

tween the quality of internal controls and the performance fees rewarded to managers, which is consistent with investors protecting against potential financial misstatements by placing less emphasis on the reported performance when internal controls are less likely to detect or prevent managers from manipulating reported performance. Finally, we find that funds in which the manager is removed from setting and reporting the official NAV to investors are significantly less likely to be subject to future regulatory investigations for fraud and/or financial misstatements.

REFERENCES

American Bar Association (ABA). 2005. *U.S. Regulation of Hedge Funds*. San Francisco, CA: Shartsis Friese LLP.

Aragon, G. O. 2007. Share restrictions and asset pricing: Evidence from the hedge fund industry. *Journal of Financial Economics* 83: 33–58.

Ashbaugh-Skaife, H., D. W. Collins, and W. R. Kinney, Jr. 2007. The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics* 44: 166–192.

———, ———, ———, and R. LaFond. 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity capital. *Journal of Accounting Research* 47: 1–43.

Blackwell, D. W., T. R. Noland, and D. B. Winters. 1998. The value of auditor assurance: Evidence from loan pricing. *Journal of Accounting Research* 36: 57–70.

Blum, J. A., M. Levi, R. T. Naylor, and P. Williams. 1998. *Financial Havens, Banking Secrecy and Money Laundering*. New York, NY: United Nations Office for Drug Control and Crime Prevention.

Brown, S. J., T. L. Fraser, and B. Liang. 2008. Hedge fund due diligence: A source of alpha in a hedge fund portfolio strategy. Working paper, New York University, New York, NY.

———, and W. N. Goetzmann. 2003. Hedge funds with style. *Journal of Portfolio Management* 26: 101–112.

———, ———, and R. G. Ibbotson. 1999. Offshore hedge fund: Survival and performance, 1989–95. *The Journal of Business* 72: 91–117.

Chan, N., M. Getmansky, S. M. Haas, and A. W. Lo. 2006. Do hedge funds increase systematic risk? *Economic Review of the Federal Reserve Board of Atlanta* 91: 49–80.

Chow, C. W. 1982. The demand for external auditing: Size, debt and ownership influences. *The Accounting Review* 42: 272–291.

Dechow, P. M., A. P. Hutton, L. Meulbroek, and R. G. Sloan. 2001. Short-sellers, fundamental analysis, and stock returns. *Journal of Financial Economics* 61: 77–106.

Doyle, J., W. Ge, and S. McVay. 2007. Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics* 44: 193–223.

Edwards, F. R., and M. O. Caglayan. 2001. Hedge fund performance and manager skill. *Journal of Futures Markets* 21: 1003–1028.

Fortin, S., and J. A. Pittman. 2007. The role of auditor choice in debt pricing in private firms. *Contemporary Accounting Research* 24: 859–896.

Fung, W., and D. A. Hsieh. 1997. Survivorship bias and investment style in the returns of CTAs. *Journal of Portfolio Management* 23: 30–41.

———, and ———. 1999. A primer on hedge funds. *Journal of Empirical Finance* 6: 309–331.

———, ———, N. Y. Naik, and T. Ramadorai. 2008. Hedge funds: Performance, risk, and capital formation. *The Journal of Finance* 63: 1777–1803.

Ge, W., and S. McVay. 2005. The disclosure of material weaknesses in internal control after the Sarbanes-Oxley Act. *Accounting Horizons* 19: 137–158.

Geczy, C. C., and J. Yan. 2006. Who are the beneficiaries when insiders trade? An examination of piggybacking in the brokerage industry. Working paper, University of Pennsylvania.

Getmansky, M., A. Lo, and I. Makarov. 2004. An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics* 74: 529–609.

Hammersley, J. S., L. A. Myers, and C. Shakespeare. 2008. Market reactions to the disclosure of internal control weaknesses and to the characteristics of those weaknesses under section 302 of the Sarbanes

Oxley Act of 2002. *Review of Accounting Studies* 13: 141–165.

HedgeFund.net (HFN). 2008. *Q1: 2008 Hedge Fund Asset Flows & Trends Report.* Available at: http://findarticles.com/p/articles/mi_m0EIN/is_2008_May_21/ai_n25455521/.

Hogan, C. E. 1997. Costs and benefits of audit quality in the IPO market: A self-selection analysis. *The Accounting Review* 72: 67–86.

Hong, H., and J. D. Kubik. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance* 58: 313–351.

Ivashina, V., and Z. Sun. 2007. Institutional stock trading on loan market information. Working paper, Harvard Business School, Cambridge, MA.

Kim, J. B., B. Y. Song, and L. Zhang. 2008. Internal control effectiveness and bank loan pricing. Working paper, Concordia University, Montreal, Canada.

Klein, B., and K. Leffler. 1981. The role of market forces in assuring contractual performance. *The Journal of Political Economy* 89: 615–638.

Krauth, B. 2007. Bounding a linear causal effect using relative correlation restrictions. Working paper, Simon Fraser University, Vancouver, Canada.

Kundro, C., and S. Feffer. 2003. *Understanding and Mitigating Operational Risk in Hedge Fund Investments.* New York, NY: The Capital Markets Company Ltd.

Lhabitant, F. S. 2008. *Handbook of Hedge Funds.* West Sussex, U.K.: Wiley Finance.

Lo, A. 2008. *Hedge Funds: An Analytic Perspective.* Princeton, NJ: Princeton University Press.

——— 2001. Risk management for hedge funds: Introduction and overview. *Financial Analysts Journal* 57: 16–33.

Managed Funds Association (MFA). 2005. *MFA's 2005 Sound Practices for Hedge Fund Managers.* Washington, D.C.: MFA.

Mansi, S. A., W. F. Maxwell, and D. P. Miller. 2004. Does auditor quality and tenure matter to investors? Evidence from the bond market. *Journal of Accounting Research* 42: 755–793.

McVea, H. 2008. Hedge fund asset valuations and the work of the International Organization of Securities Commissions (IOSCO). *The International and Comparative Law Quarterly* 57: 1–24.

Oesterle, D. A. 2006. Regulating hedge funds. *Entrepeneurial Business Law Journal* 1: 1–42.

Ogneva, M., K. Raghunandan, and K. R. Subramanyam. 2007. Internal control weakness and cost of equity: Evidence from SOX Section 404 disclosures. *The Accounting Review* 82: 1255–1297.

PricewaterhouseCoopers. 2006. *The Regulation, Taxation and Distribution of Hedge Funds in Europe: Changes and Challenges.* New York, NY: PwC.

Rosenbaum, P. 2002. *Observational Studies.* New York, NY: Springer.

Securities and Exchange Commission (SEC). 2003. *Implications of the Growth of Hedge Funds.* Staff Report. Washington, D.C.: SEC.

Smith, T. R., Jr. 2006a. Mutual funds under fire: A chronology of developments since January 1, 2003. *Journal of Investment Compliance* 7: 4–33.

———. 2006b. Mutual funds under fire: reform initiatives. *Journal of Investment Compliance* 7: 4–27.

Stickel, S. E. 1992. Reputation and performance among security analysts. *The Journal of Finance* 47: 1811–1836.

Suss, E. C., O. H. Williams, and C. Mendis. 2002. Caribbean offshore financial centers: Past, present, and possibilities for the future. Working paper, International Monetary Fund.

Watts, R. L., and J. L. Zimmerman. 1983. Agency problems, auditing, and the theory of the firm: Some evidence. *The Journal of Law & Economics* 23: 613–633.

Weber, J., M. Willenborg, and J. Zhang. 2008. Does auditor reputation matter? The case of KPMG Germany and ComROAD AG. *Journal of Accounting Research* 46: 941–972.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 1921–1949

American Accounting Association
DOI: 10.2308/accr.2010.85.6.1921

Discretion in Bonus Plans

Merle Ederhof
University of Michigan

ABSTRACT: This study examines discretionary bonus payments by firms to senior-level executives. Interpreting discretionary bonuses as the result of implicit incentive contracts, I analyze an analytical model that includes a contractible and a non-contractible performance measure. The model yields the primary hypothesis that discretionary bonuses occur when the outcome of the contractible measure is either low or high, but not when the contractible outcome falls in the medium range. Based on a sample collected from public sources, I find empirical support for the notion that discretionary bonuses are paid based on non-contractible performance measures that are related to future financial performance. Moreover, discretionary bonus payments occur significantly more often when the contractible performance measure falls in the tails of the distribution. In contrast, I do not find support for the predictions that discretionary bonus payments are related to the manipulability of the contractible performance measures or that discretionary bonus payments are related to the power of the executives in the companies.

Keywords: *discretionary bonus; incentive compensation; implicit contract; non-contractible performance measure; managerial power.*

Data Availability: *All data used in this study are publicly available.*

I. INTRODUCTION

This study examines discretionary bonus payments. Discretionary bonus payments are bonus payments that are not determined by a strict formulaic approach. For example, Sun Microsystems' chief executive officer Scott McNealy did not receive a bonus according to his bonus formula for fiscal year 2005; however, he was paid a $1.1 million discretionary bonus. This study addresses the question of why companies pay discretionary bonus payments to their executives.

First, I investigate when discretionary payments may occur in an optimal incentive contract using an analytical model. I hypothesize that discretionary bonuses reflect the board's assessment of non-contractible performance measures. Therefore, my model includes not only a contractible

This study is based on my dissertation at the Stanford Graduate School of Business. I thank the members of my dissertation committee: David Larcker (Chair), Madhav Rajan, and Stefan Reichelstein. I also thank Rick Antle, Anne Beyer, Harry Evans III (editor), Brian Mittendorf, Jake Thomas, two anonymous reviewers, and seminar participants at Stanford University, University of California, Davis, The University of Utah, University of Washington, Massachusetts Institute of Technology, University of Rochester, The University of Chicago, Harvard University, New York University, Emory University, Yale University, and University of Pennsylvania.

Editor's note: Accepted by John H. (Harry) Evans III.

Submitted: May 2008
Accepted: March 2010
Published Online: November 2010

performance measure, but also a non-contractible performance measure.[1] The model is based in a conventional moral hazard setting but differs from traditional agency models, which only include contractible performance measures. The model developed for this study extends Rajan and Reichelstein (2009), yielding the prediction that discretionary bonus payments occur when the outcome of the contractible performance measure is either low or high, but not when the outcome of the contractible performance measure falls in the medium range.

Baker et al. (1994) also develop a model that includes a contractible and a non-contractible performance measure. In contrast to the one-period model developed in this study, the model in Baker et al. (1994) is based in an infinite horizon setting. Their model yields the prediction that discretionary bonuses are more likely to be used when the contractible measure is more susceptible to manipulation.

The predictions from the analytical models are contrasted in the empirical analysis with a prediction based on managerial power theory. Managerial power theory predicts a positive relationship between discretionary bonus payments and the power of the executives in the company. Thus, managerial power theory suggests that discretionary bonus payments are more likely to occur in companies with weak corporate governance structures. According to managerial power theory, executives who have considerable power over the board of directors and the compensation committee receive compensation above the level explained by standard economic factors alone (e.g., Lambert et al. 1993; Boyd 1994; Core et al. 1999; Bebchuk et al. 2008). Discretionary bonus payments may be one way in which powerful executives achieve such excess pay (Bebchuk et al. 2002).

I test the hypotheses using a sample collected from public sources. The data are gathered from companies' Forms 8-K and proxy statements starting in August 2004 when the updated SEC regulation regarding the disclosure of discretionary bonus payments in Forms 8-K became effective. This regulation requires companies to disclose the amounts of discretionary bonuses paid to the top five executives. Thus, the sample includes explicit information on the amount of discretionary bonus payments over the period from August 2004 to September 2006, and consists of 234 firm-years.

The results of the analyses provide support for the hypothesis that discretionary bonuses reflect implicit incentive contracts based on non-contractible performance measures. Specifically, discretionary bonuses are incrementally informative about future financial performance after controlling for current period financial performance, indicating that discretionary bonuses are paid based on forward-looking performance measures that are only observable to the contracting parties and that are likely to be non-contractible. The results also support the prediction derived from the model that companies are more likely to pay discretionary bonuses when the contractible outcome falls in the tails of the distribution. Conversely, the results do not support the hypotheses that there is a positive relationship between the occurrence of discretionary bonuses and the manipulability of the contractible measures or the power of the executives in the companies. Overall, the results support the notion that companies use discretionary bonus payments in a way that is consistent with optimal incentive contracting.

The paper contributes to the literature on discretion in bonus plans in several ways. First, I provide evidence that companies are more likely to pay discretionary bonuses when the contract-

[1] Examples of contractible performance measures that are frequently used are accounting-based measures such as earnings, revenue, or cash flow. Examples of non-contractible performance measures include subjective assessments of the agent's performance by the principal: whether the agent negotiated new deals with major customers or suppliers that are likely to pay off in the future; whether the agent implemented important strategic initiatives; whether the agent worked well as a team member; whether the agent was a good leader; assessment of the agent's initiative, or whether the agent trained subordinates well.

ible performance is either very low or very high. The model I develop shows that this finding is consistent with optimal incentive contracting. Second, this study is one of the first to link discretionary bonus payments to corporate governance. Although shareholder activists and the press have portrayed discretionary bonus payments as a form of rent extraction, researchers have not examined the relationship between discretionary bonus payments and executive power. Finally, the sample used in this study improves upon existing empirical work by using an explicit measure of discretion for companies in a wide array of industries.[2]

The next section develops the analytical model and the hypotheses. Section III describes the sample and measures used in the empirical analysis. Section IV discusses the research design and the empirical results. Section V provides a summary and conclusion.

II. ANALYTICAL MODEL AND HYPOTHESES DEVELOPMENT

Economic Theory

Analytical research has modeled discretionary bonus payments as the result of implicit incentive contracts based on non-contractible performance measures (e.g., Baiman and Rajan 1995).[3] The model below follows that approach. In order to link the empirical analysis to the theoretical work, it is desirable to provide support for the modeling approach of capturing discretionary bonus payments as the result of non-contractible performance measures. Since implicit incentive contracts used in firms are unobservable, I cannot directly test whether discretionary bonus payments are indeed paid based on non-contractible performance measures.

However, Hayes and Schaefer (2000) have introduced an approach to indirectly test whether the performance metrics underlying the bonus schemes are likely to be non-contractible in nature. In particular, Hayes and Schaefer (2000) regress future financial performance on current cash compensation and on current financial performance. They find that current compensation is incrementally informative about future financial performance after controlling for current financial performance. Their findings are consistent with the notion that companies employ performance metrics that are only observable to the contracting parties and that are correlated with future financial performance.

Hayes and Schaefer (2000) estimate their regression using the sum of salary and bonus as the measure of current compensation. My data set allows a separation of the bonus payment into its discretionary and formula-based components. Since the formula-based portion of the bonus payment is expected to be based on performance metrics such as accounting earnings, I expect that the formula-based bonus component is not incrementally informative about future financial performance. In contrast, I hypothesize that the discretionary bonus components are based on non-contractible performance measures, which are likely to be correlated with future performance and unobservable to outside parties. In other words, I hypothesize that the discretionary bonus components are incrementally informative about future financial performance. These arguments lead to my first hypothesis. (All hypotheses are stated in the alternative form).

H1: The discretionary components of bonus payments are incrementally informative about future financial performance, whereas the formula-based components are not.

In order to derive predictions regarding the occurrence of discretionary bonus payments under optimal incentive contracting, I interpret discretionary bonuses as the result of implicit incentive contracts based on non-contractible performance measures. By definition, non-contractible perfor-

---

[2] Prior empirical studies that have analyzed discretion in bonus plans include Murphy and Oyer (2003) and Gibbs et al. (2004).

[3] Implicit contracts are agreements that are not legally enforceable. They are sustained through economic mechanisms and constitute equilibria in a game-theoretic sense (Bull 1987).

mance measures are unverifiable, such that it is impossible to write explicit incentive schemes based on such metrics. Thus, in order to credibly incorporate non-contractible information, the principal has to resort to an economic enforcement mechanism.

One such mechanism is for the principal to commit to paying out a fixed bonus pool that is independent of the outcome of the non-contractible performance measure (MacLeod 2003; Rajan and Reichelstein 2006; Rajan and Reichelstein 2009). The implicit agreement between the contracting parties specifies how the bonus pool will be divided between the agent and a third party, depending on the realization of the non-contractible signal.[4] Since the principal is *ex post* indifferent as to the recipient of the bonus pool, the presumption is that she will follow the implicit contract that was agreed upon *ex ante*.

Rajan and Reichelstein (2009; hereafter RR) also analyze a setting where, in addition to the non-contractible performance measure, the principal also has access to a contractible signal. This setting is particularly relevant for my empirical setting since, aside from the discretionary payments, the companies that comprise my sample also have formal bonus plans in place that are largely based on financial, and therefore contractible, performance measures.

The following analysis builds on the model analyzed in RR. In particular, the principal seeks to implement a given action, a^h, which is more costly for the agent than a less productive action, a^l. The principal can use both a contractible performance measure, x, and a non-contractible performance measure, y, in designing the contract. Both performance measures are informative about the agent's action and the outcomes of both measures are binary, i.e., $x \in \{x^l, x^h\}$ and $y \in \{y^l, y^h\}$. Moreover, the monotone likelihood ratio property (MLRP) is assumed to hold for both performance measures.[5]

The analysis in RR is restricted to the case for which the two performance measures are independent (conditional on a given action). In contrast, I consider the more general setting where the two metrics may be conditionally dependent. For example, in many scenarios it seems reasonable to expect the contractible and non-contractible measures to be positively correlated. In particular, a positive macroeconomic shock may result in a strong accounting performance and may also lead to a favorable subjective assessment of the agent's performance since factors such as strategic initiatives and new customers are likely to be influenced by the overall economic environment as well.

Thus, following RR, the contractible measure reflects the agent's action in the following way:

$$p_h^h \equiv \text{Prob}[x = x^h | a = a^h] > p_h^l \equiv \text{Prob}[x = x^h | a = a^l].$$

In contrast to the setup in RR, the probability of observing a high outcome on the non-contractible measure, given that the agent took action a^i, may depend on the realization of the contractible measure x. In particular:

$$q_{jh}^i \equiv \text{Prob}[y = y^h | a = a^i, x = x^j],$$

where $i, j \in \{l, h\}$.

The remaining structural elements follow the setup in RR. In particular, the bonus pool corresponding to the contractible outcome j is denoted by w^j. The payment that the agent receives for contractible outcome j and non-contractible outcome k is denoted by s^{jk}, which satisfies the inequality $w^j \geq \max\{s^{jh}, s^{jl}\}$. Any remaining difference between the bonus pool and the payment

---

[4] One interpretation of the third party is a charity. More generally, the third party can be thought of as any agent (inside or outside the organization) that is not involved in the contracting relationship.

[5] In the context of my model, the density $p_j(a)$ satisfies the MLRP if $\frac{p_j(a^l)}{p_j(a^h)}$ is monotone decreasing in j. The MLRP implies that higher outcomes of the performance measure are indicative of higher effort on the part of the agent.

to the agent, $(w^j - s^{jk})$, is diverted to the third party. Moreover, the risk-averse agent is assumed to have additively separable preferences over wealth, $U(\cdot)$ and cost of effort, $e(\cdot)$ and the principal's optimization problem can be formulated as follows:

Program P1:

$$\underset{w^j, s^{jk}}{\text{Min}}[w^h \cdot p_h^h + w^l \cdot (1 - p_h^h)]$$

subject to:
(i) $E[U(s^{jk}) \mid a^h] - e(a^h) \geq \bar{U}$,
(ii) $E[U(s^{jk}) \mid a^h] - e(a^h) \geq E[U(s^{jk}) \mid a^l] - e(a^l)$,
(iii) for all j, k: $w^j - s^{jk} \geq 0$.

The non-contractible performance measure is said to be used if the payment to the agent depends on the outcome of y. Specifically, the non-contractible measure is said to be used for contractible outcome $x = x^j$ if the solution to P1 is such that $Min(s^{\hat{j}h}, s^{\hat{j}l}) < \hat{w}^j$; the non-contractible measure is said to not be used for contractible outcome $x = x^j$ if the solution to P1 is such that $Min(s^{\hat{j}h}, s^{\hat{j}l}) = \hat{w}^j$.

RR identify necessary and sufficient conditions for the non-contractible measure to optimally be used for the low-contractible outcome (RR, Proposition 1 and Observation 1). Since the setting in RR is a special case of the setting analyzed here (i.e., $q_{lh}^i = q_{hh}^i = q_h^i$), it follows directly that there exist circumstances where the optimal contract for the present setup is such that the non-contractible measure is used for the low-contractible outcome.

In contrast to the findings in RR, the following result shows that, if the assumption of conditional independence is relaxed, it is optimal for the principal to employ the non-contractible indicator for the high-contractible outcome in certain circumstances. Let $Q^L \equiv \frac{1-q_{lh}^l}{1-q_{lh}^h}$ and $Q^H \equiv \frac{1-q_{hh}^l}{1-q_{hh}^h}$ where $Q^L, Q^H \in (1, \infty)$. Q^L and Q^H reflect the informativeness of the non-contractible performance measure for the low- and high-contractible outcome, respectively. I also use the notation:

$$P^* = \frac{1 - T \cdot \dfrac{p_h^l \cdot (1 - p_h^h)}{p_h^h \cdot (1 - p_h^l)}}{1 - T}$$

where:

$$T = \frac{V\left(\bar{U} + \dfrac{e(a^l) \cdot p_h^h - e(a^h) \cdot p_h^l}{(p^h - p^l)}\right)}{V\left(\bar{U} + \dfrac{e(a^h) \cdot (1 - p_h^l) - e(a^l) \cdot (1 - p_h^h)}{(p_h^h - p_h^l)}\right)}$$

and $V(\cdot)$ represents the derivative of the inverse of the agent's utility function, $U^{-1}(\cdot)$. The result identifies sufficient conditions for the non-contractible performance measure to optimally be used for the high outcome of x. All proofs are in Appendix A.

Proposition 1: The non-contractible performance measure, y, is used when the outcome of the contractible performance measure, x, is high as long as the following conditions are satisfied:

$$Q^H > P^* \cdot \left[\frac{p_h^h \cdot (1-p_h^l)}{p_h^l \cdot (1-p_h^h)}\right] > Q^L \cdot \left[\frac{p_h^h \cdot (1-p_h^l)}{p_h^l \cdot (1-p_h^h)}\right]. \quad (1)$$

Following RR, the intuition for this result is that using the non-contractible performance measure in the incentive scheme entails an additional agency cost for the principal since she has to divert money to a third party with positive probability. Thus, the principal will only use the non-contractible metric if doing so is sufficiently gainful and she will use it in a way that is most beneficial. When the conditions in (1) are satisfied, the non-contractible signal for the high-contractible outcome is sufficiently informative such that the principal finds it beneficial to use the non-contractible metric for the high-contractible outcome.

Having established that the non-contractible performance measure may optimally be used for either the low or the high outcome of the contractible measure in the binary outcome case, to determine whether the non-contractible signal would ever be used for intermediate values of the contractible measure, I next allow three possible contractible performance measure outcomes.[6] The setup of the model is the same as before, except now $x \in \{x^l, x^m, x^h\}$.[7] The distributions of x are now given by:

$$p_j^i \equiv \text{Prob}[x = x^j | a = a^i]$$

where $i \in \{l,h\}$ and $j \in \{l,m,h\}$. Analogous to the case where the contractible performance measure has two outcomes, the probability that the outcome of the non-contractible performance measure is high is denoted by q_{jh}^i where, now, $i \in \{l,h\}$ and $j \in \{l,m,h\}$. The principal's optimization problem now becomes:

Program P2:

$$\underset{w^j, s^{jk}}{\text{Min}}[w^h \cdot p_h^h + w^m \cdot p_m^h + w^l \cdot (1 - p_h^h - p_m^h)]$$

subject to:
(i) $E[U(s^{jk}) \mid a^h] - e(a^h) \geq \bar{U}$,
(ii) $E[U(s^{jk}) \mid a^h] - e(a^h) \geq E[U(s^{jk}) \mid a^l] - e(a^l)$,
(iii) for all j, k: $w^j - s^{jk} \geq 0$ (six constraints).

Proposition 2 characterizes sufficient conditions for the non-contractible metric not to be used when the contractible measure takes on the intermediate value. To represent the informativeness of the non-contractible performance measure for the intermediate outcome of the contractible measure, let $Q^M \equiv \frac{1-q_{mh}^l}{1-q_{mh}^h}$.

Proposition 2: When the contractible performance measure, x, has three possible outcomes, the non-contractible performance measure, y, is not used for the intermediate outcome of the contractible performance measure if $Q^L \geq Q^M$.

As an alternative to the condition in Proposition 2, consider the following case. It seems plausible that the probability that the principal infers a high subjective performance may be increasing in the outcome of the contractible measure in the sense that $q_{hh}^h \geq q_{mh}^h \geq q_{lh}^h$ and $q_{hh}^l \geq q_{mh}^l \geq q_{lh}^l$. These conditions imply that x and y are positively correlated. As discussed earlier, a

[6] I restrict attention to the case where the contractible performance measure has three outcomes since the ultimate goal is to derive predictions for the empirical analysis. The case where the contractible performance measure has n outcomes adds little insight for the empirical predictions and would complicate the analysis.

[7] In particular, the agent's action and the outcome of the subjective performance measure are still binary and the MLRP is again assumed for both performance measures.

positive macroeconomic shock could yield both a strong accounting performance and a favorable subjective assessment of the agent's performance.

A second interpretation of the conditions $q^h_{hh} \geq q^h_{mh} \geq q^h_{lh}$ and $q^l_{hh} \geq q^l_{mh} \geq q^l_{lh}$ is that the subjective performance evaluation technology is better at reflecting the agent's action when the contractible outcome more accurately reflects the action taken by the agent. Stated differently, $q^h_{hh} \geq q^h_{mh} \geq q^h_{lh}$ indicates that the degree to which the subjective performance evaluation technology (correctly) yields a favorable assessment of the agent's performance is increasing in the contractible outcome. Analogously, $q^l_{hh} \geq q^l_{mh} \geq q^l_{lh}$ indicates that the degree to which the subjective performance evaluation technology (correctly) yields an unfavorable assessment of the agent's performance is higher when the contractible outcome is lower.

The following analysis shows that if we make the additional assumption that the improvement in the subjective performance evaluation technology described above is non-decreasing, i.e., $(q^h_{hh} - q^h_{mh}) \geq (q^h_{mh} - q^h_{lh})$ and $(q^l_{hh} - q^l_{mh}) \leq (q^l_{mh} - q^l_{lh})$, then we also obtain the result that the non-contractible measure is not used for intermediate values of the contractible metric.[8] In particular, Appendix B summarizes the results of Monte Carlo simulations that have been conducted in order to characterize the optimal incentive contract in this setting. In running the simulations, I follow the prior literature (e.g., Baiman et al. 1995; Baiman and Rajan 1995; Rajan and Reichelstein 2006, 2009) in choosing the form of the agent's utility function and the parameter spaces for $e(a^l)$, $e(a^h)$, and $\bar{U}$. Specifically, I assume that the agent has a power utility function of the form $U(s) = \frac{1}{1-\gamma} s^{(1-\gamma)}$. The simulation results indicate that, for the chosen parameter spaces, the non-contractible signal is not used when the contractible signal takes on the intermediate value.[9] The simulations also indicate that the situations where the non-contractible measure is used for the low-contractible outcome are more frequent than the situations where the non-contractible signal is used for the high-contractible outcome.

The result in Proposition 2 and the simulation analysis forms the basis for the next prediction:

H2: Discretionary bonus payments are more likely to occur when the outcome of the contractible performance measure is either low or high than when the outcome of the contractible performance measure falls in the medium range.

The model described above focuses on the principal's committing to a bonus pool as the mechanism that enforces the implicit contract. Furthermore, prior literature has proposed alternative mechanisms that can be employed to credibly incorporate subjective information into incentive contracts. In particular, Baker et al. (1994) analyze the interplay of contractible and non-contractible performance measures in a setting with repeated interaction between the principal and the agent, where the implicit contract is enforced by the principal's concern for her reputation.[10] In their model, Baker et al. (1994) focus on trigger strategies as the equilibrium concept, which stipulates that the principal can never again employ the non-contractible metric once she has reneged on the implicit contract. Thus, in deciding whether to honor the implicit contract, the principal compares the firm's future profits when both measures can be used for contracting to the

---

[8] The conditions $q^h_{hh} \geq q^h_{mh} \geq q^h_{lh}$, $q^l_{hh} \geq q^l_{mh} \geq q^l_{lh}$, $(q^h_{hh} - q^h_{mh}) \geq (q^h_{mh} - q^h_{lh})$, and $(q^l_{hh} - q^l_{mh}) \leq (q^l_{mh} - q^l_{lh})$ (which are statements for a given action taken by the agent) together imply that the informativeness of the non-contractible measure (which is defined across the two possible actions) is not highest for the intermediate-contractible outcome. Stated differently, the conditions imply that either $Q^L \geq Q^M$ and/or $Q^H \geq Q^M$. Intuitively, the subjective performance evaluation technology is so accurate at identifying the low action for the low-contractible outcome and at identifying the high action for the high-contractible outcome that the informativeness of the subjective signal will be highest in one or both of the tails.

[9] In future research, it would be desirable to explore whether this result can also be shown analytically.

[10] Other papers that analyze implicit contracts in infinite horizon settings include Bull (1987), Pearce and Stacchetti (1998), Levin (2003), and Budde (2007).

profits that can be generated when only the contractible measure remains credible. The authors show that, when the contractible performance measure is less "distorted," this difference in future profits is smaller, and the non-contractible measure becomes less attractive for contracting purposes. Therefore, the optimal incentive contract is such that the weight placed on the non-contractible measure is increasing in the "distortion" of the contractible measure.[11]

Baker et al. (1994) capture the quality of the contractible performance measure by the variance of the marginal product, which maps the agent's action into the performance measure outcome. In the model, the agent receives private information regarding the marginal product before choosing his action. The authors interpret a measure with a higher-variance marginal product as a performance measure that is more "distorted" and that is associated with a higher level of dysfunctional behavior. Along those lines, prior empirical work has interpreted the variance of the marginal product in the Baker et al. (1994) model as a performance measure's susceptibility to manipulation (Gibbs et al. 2004), and the potential for earnings management (Bushman et al. 1996). For the purpose of this study, I follow these existing empirical studies and, given that the vast majority of companies use a measure of accounting profits in their bonus plans (e.g., Murphy 2000), interpret the distortion of the contractible metric as the manipulability of the company's earnings. This leads to the following hypothesis:

H3: Discretionary bonus payments are more likely to occur when the company's earnings are more susceptible to manipulation.

Managerial Power Theory

The preceding hypotheses are based on the notion that discretionary bonus payments are an outcome of optimal incentive contracts. Alternatively, discretionary bonus payments may constitute managerial rent extraction stemming from governance limitations in the companies. In other words, discretionary bonus payments could conceivably occur in settings where management has considerable power over the board of directors and the compensation committee.

Prior empirical studies have documented that measures of CEO power are associated with several compensation patterns that can be interpreted as rent extraction. For example, Core et al. (1999) document that CEO power is positively related to excessive pay levels (i.e., compensation above the level explained by standard economic factors). Moreover, Bebchuk et al. (2008) show that powerful CEOs are more likely to receive a "lucky" option grant at the lowest price of the month and that they are more likely to be rewarded for luck due to industry-wide shocks.

Discretionary bonus payments may be one way in which powerful executives extract rents from their companies. Bebchuk et al. (2002) argue that an important factor in executives' ability to extract rents is the amount of "outrage" that a proposed pay package would create. If the outrage over a pay package is sufficiently high and widespread, then it imposes significant costs, which limit the amount of rents that the executives can extract. Bebchuk et al. (2002) argue further that the outrage costs depend on the extent to which the rent extraction can be easily and distinctly identified. The authors argue that, under the managerial power approach, executives prefer compensation structures that enable the extraction of rents to be camouflaged as optimal contracting.

As suggested in the preceding section, under certain conditions, discretionary bonus payments will be part of the optimal incentive contract. Moreover, the incentive plans that shareholders

[11] Baker et al. (1994) obtain the result that the weight placed on the non-contractible performance measure *decreases* with the distortion of the contractible measure in the case that the firm has to shut down after the firm reneges on the implicit contract. However, this case is unlikely to reflect my empirical setting since I would expect that companies can still generate positive returns when restricted to using contractible measures such as accounting earnings and the company's stock price in incentive contracting.

approve typically allow for the payment of discretionary bonuses. Stated differently, companies can usually pay discretionary bonuses without an additional approval of shareholders. Therefore, it is difficult for outsiders to determine whether a discretionary bonus payment is paid in accordance with the optimal incentive contract or whether the discretionary payment constitutes rent extraction. Following the arguments provided by Bebchuk et al. (2002), discretionary bonus payments may be an effective way for powerful executives to extract rents. Hypothesis 4 follows from this argument:

H4: Discretionary bonus payments are more likely to occur in companies where management has considerable power.

III. SAMPLES AND MEASURES

I define "discretionary bonuses" as payments that would not have been paid according to the explicit contract entered into before the agent's actions are taken. This definition includes bonuses that are paid according to implicit incentive contracts based on non-contractible performance measures. The definition excludes bonuses that are paid according to explicit incentive contracts based on contractible performance measures. In addition, the definition excludes payments such as signing bonuses, severance payments, and guaranteed minimum payments that are based on explicit contractual agreements.

Sample

The sample used in this study is collected from public sources and consists of two subsamples: (1) a sample of companies that paid discretionary bonuses ("discretionary sample"), and (2) a control sample that did not make these types of payments to executives ("non-discretionary sample").

The discretionary sample is collected by taking advantage of the updated SEC regulation regarding disclosure requirements in Forms 8-K that became effective in August 2004. In particular, companies are required to disclose payments of material cash bonuses if the bonuses are not paid in accordance with the bonus formula.[12] Since companies may disclose the required information in a subsequent proxy statement if they fail to file a Form 8-K, the discretionary sample consists of companies that disclose discretionary bonus payments in either a Form 8-K or a proxy statement between August 23, 2004 and September 20, 2006.[13] The sample is constructed based on a keyword search of Forms 8-K and proxy statements contained in the database "Lexis/Nexis Research" using several different search strings.[14] To be included in the discretionary sample, the company must disclose a dollar amount for the discretionary bonus payment in the Form 8-K or proxy statement. The final discretionary sample consists of 100 firm-years from the search of Forms 8-K and 134 firm-years from the search of proxy statements for a total of 234 firm-years and 211 firms.

The non-discretionary sample consists of companies that do not disclose information about discretionary bonus payments in either a Form 8-K or a proxy statement over the period between August 23, 2004 and September 20, 2006. Since some of the tests require information on the target bonus, the non-discretionary sample is constructed by applying several search strings related to

---

[12] More precisely, according to a statement by the Division of Corporation Finance of the SEC dated November 23, 2004: "Payment of a cash bonus must be disclosed under Item 1.01 of Form 8K if the registrant exercised discretion to pay the bonus even though the specified performance criteria were not satisfied."

[13] More precisely, if companies fail to timely file a Form 8-K, a safe harbor rule applies as long as the information is disclosed in a Form 10-Q or a Form 10-K relating to the period in which the event occurred.

[14] I searched the Forms 8-K and proxy statements using keywords such as "discretionary bonus," "special bonus," "additional bonus," "discretionary award," "discretionary cash award," etc.

target bonuses to the proxy statements contained in "Lexis/Nexis Research" that were *not* identified by Lexis/Nexis when constructing the discretionary sample.[15] To be included in the non-discretionary sample, the company must disclose the level of either the target bonus or the maximum bonus. Moreover, in order to minimize the influence of signing bonuses, documents where the managers had a different position in the year before are also excluded from the sample. For a subset of 50 companies in the non-discretionary sample, I extensively searched the proxy statement and the Forms 8-K for information regarding discretionary bonuses. Using several broad keyword searches, I could not find any information regarding the payment of discretionary bonuses. The final non-discretionary sample consists of 349 firm-years and 279 firms. The final total sample consists of 583 firm-years and 490 firms. Table 1 provides further details on the sample collection process.

Measures

Bonus Payment Components

Total bonus payment is measured by the sum of the formula-based bonus and the discretionary bonus, scaled by salary (*TOTALBONUS*), and its formula-based and discretionary components are denoted by (*FORMULABONUS*) and (*DISCBONUS*). Tests that concern the occurrence of discretionary bonus payments are estimated using an indicator variable that is set equal to 1 if a discretionary bonus was paid, and 0 otherwise (*DISCIND*).

Financial Performance

The first hypothesis predicts that discretionary bonus payments are predictive of future financial performance after controlling for current financial performance. Similar to the measures used in Hayes and Schaefer (2000), I use the firm's return on assets in year t+1 (ROA_{t+1}) as the measure of future financial performance. To measure the firm's current performance, I employ the current year's return on assets (ROA_t), stock return (RET_t), and log of sales (log $SALES_t$).

Outcome of Contractible Performance Measures

Hypothesis 2 requires a measure of the outcome(s) of the contractible performance measure(s). I measure this construct in two ways: (1) using the bonus payment that was made according to the bonus formula as a percentage of the target bonus (*FBONUS*), and (2) using the change in earnings-per-share (ΔEPS).

Using the formula-based bonus payment seems natural since there is a one-to-one mapping between the outcome(s) of the performance measure(s) used in the bonus formula and the bonus payment based on the outcome(s). A limitation of the measure *FBONUS* is that its value is influenced by how the target bonus is set. Prior empirical studies have provided evidence suggesting that the performance target corresponding to the target bonus may not necessarily reflect an unbiased performance expectation. In particular, using various settings and samples, studies have documented that (1) actual performance exceeds targeted performance in the majority of cases (Merchant and Manzoni 1989; Leone and Rock 2002; Indjejikian and Nanda 2002), (2) performance target revisions follow certain patterns (Leone and Rock 2002; Indjejikian and Nanda 2002; Indjejikian and Matejka 2006), and (3) performance targets may be based on peer groups, timeless standards, or the company's cost of capital (Murphy 2000). Because of this limitation, I follow

[15] I used keywords such as "target bonus" or "maximum bonus."

TABLE 1

Sample Collection

I. Companies that Paid Discretionary Bonuses ("discretionary sample")	
I. a. Forms 8-K	
Documents identified by Lexis/Nexis	1164
No discretionary bonus was paid	854
Employment agreement, separation agreement, promotion agreement	83
Not clear whether bonus is discretionary	88
Proxy statement not available	32
Not a named officer	7
Final Firm-Years	100
I. b. Proxy Statements	
Documents identified by Lexis/Nexis	854
No discretionary bonus was paid	472
Employment agreement, separation agreement, promotion agreement	56
Not clear whether bonus is discretionary	137
Already in 8-K Sample	55
Final Firm-Years	134
I. c. Final Discretionary Sample	
Final Firm-Years	234
Final Firms	211
II. Companies That Did Not Pay Any Discretionary Bonuses ("non-discretionary sample")	
Documents identified by Lexis/Nexis	952
Amount of target/maximum bonus not available	470
Manager had a different position in the prior year	133
Final Firm-Years	349
Final Firms	279
III. Total Sample	
III. a. Firm-Years	
Companies that paid discretionary bonuses	234
Companies that did not pay any discretionary bonuses	349
Final Firm-Years	583
III. b. Firms	
Companies that paid discretionary bonuses	211
Companies that did not pay any discretionary bonuses	279
Final Firms	490

prior empirical studies (Sloan 1993; Core et al. 2003) and also conduct the analyses using the change in EPS.

Manipulability of Earnings

Hypothesis 3 predicts that discretionary bonuses are more likely to occur when a company's earnings are more susceptible to manipulation. I measure this construct using a variable that captures the volatility of the company's accruals. More precisely, I calculate the standard deviation of the scaled total accruals over the prior five years (*SDTA*). This measure captures the spirit of the

earlier discretionary accruals models (Healy 1985; DeAngelo 1986).[16] Moreover, accrual volatility is correlated with the measure of accrual quality that is proposed by Dechow and Dichev (2002, 48).[17] Companies with higher levels of accrual volatility are expected to have earnings that are more susceptible to manipulation.

Executive Power

Hypothesis 4 requires a measure of executive power. I employ the measure "CEO pay slice" (*CPS*), which is defined as the percentage of the total compensation paid to the top five executives that goes to the CEO. This measure, introduced by Bebchuk et al. (2008), has received increasing attention from the corporate governance literature. Bebchuk et al. (2008) show that this measure is related to the odds of awarding "lucky" option grants at the lowest price of the month and to the tendency to reward the CEO for luck due to positive industry-wide shocks. *CPS* is expected to reflect the power and influence of the CEO in the organization, meaning that CEOs with a higher *CPS* are expected to have a greater influence on the compensation-setting process. With respect to my empirical setting with a relatively small sample, this measure also has the advantage that it can easily be computed from a company's compensation table and does not require additional data sources.[18] All measures are defined and summarized in Table 2.

Descriptive Statistics

Table 3 provides descriptive statistics for the variables as they are used in the analyses. Panel A provides descriptive statistics for the discretionary subsample and Panel B provides descriptive statistics for the non-discretionary subsample.

The executives in the companies that made discretionary payments experience increases in their relative bonus payments from the prior year, as reflected in the positive mean and median of ($\Delta TOTALBONUS$). However, a closer look reveals that the median increase in the formula-based bonus component ($\Delta FORMULABONUS$) is zero. Thus, the executives receive higher total bonus payments, on average, as compared to the prior year because of the discretionary bonus component.

The descriptive statistics for the measures reflecting financial performance indicate that, in the year for which the discretionary bonuses were paid, the companies, on average, experience growth in sales and ROA but a decline in market returns. Moreover, the companies experience further growth in ROA in the year following the discretionary payments.

The descriptive statistics for *FBONUS*, which captures the ratio of (formula bonus/target bonus), indicate that the executives in the companies that comprise the discretionary sample, on average, do not achieve their targeted performance. Conversely, the executives in the companies that do not pay discretionary bonuses on average achieve a performance around the target level. The mean (median) values for *FBONUS* are 0.84 (0.83), and 1.06 (1.00) for the discretionary and non-discretionary sample, respectively. In contrast to the descriptive statistics for *FBONUS*, the descriptive statistics for the change in EPS suggest that the performance is comparable across the two samples.

---

[16] Estimating more sophisticated discretionary accrual models such as the Jones (1991) model would result in a high proportion of missing data points in my relatively small sample.

[17] Dechow and Dichev (2002) focus on working capital accruals. This study focuses on total accruals due to missing data problems in calculating working capital accruals for my relatively small sample. As discussed by Dechow and Dichev (2002), the notion of accrual quality differs from the concept of discretionary accruals. However, the "distortion" in contractible performance measures used in the model in Baker et al. (1994) could also be interpreted such that higher "distortion" is related to lower accrual or earnings quality.

[18] The results are robust to using alternative proxies for executive power that are based on things like the percentage of outside directors that were appointed by existing inside directors, the number of directors on the board, whether the company has a staggered board, etc. Using these alternative proxies would result in a smaller sample.

TABLE 2

Measures

Measures of Bonus Payment Components	
TOTALBONUS	((Formula-based bonus + discretionary bonus)/salary)
FORMULABONUS	(Formula-based bonus/salary)
DISCBONUS	(Discretionary bonus/salary)
DISCIND	Indicator variable that is equal to 1 if a discretionary bonus was paid; 0 otherwise
Measures of Financial Performance	
ROA	(Net income/average total assets)
RET	Market return
SALES	Sales
Measures of Contractible Performance	
FBONUS	(Formula bonus/target bonus); for companies that only provide information on the maximum bonus: (Formula bonus/0.5 * maximum bonus)
ΔEPS	$((EPS_t - EPS_{t-1}) / \text{Share Price}_{t-1})$
Measure of Manipulability of Earnings	
SDTA	Log of the standard deviation of (total accruals/average total assets) over the prior five years; where Total accruals = (Income before extraordinary items − operating cash flows)
Measure of Executive Power	
CPS	(Total CEO compensation)/(Aggregate total compensation for the five highest-paid executives), where: Total compensation = (Salary + bonus + other annual pay + Black-Scholes value of options granted + value of restricted stock grants + value of long-term incentive payout + all other compensation)

Contrary to the prediction in H3, the discretionary subsample does not exhibit greater accrual volatility than the non-discretionary subsample (the mean and median values for *SDTA* are −3.18 and −3.19 for the discretionary subsample and −3.23 and −3.43 for the non-discretionary subsample). Moreover, the values for the variable capturing the CEO's pay slice do not suggest significant differences with regard to executive power in the two samples.

IV. RESEARCH DESIGN AND EMPIRICAL RESULTS

Future Financial Performance

Hypothesis 1 predicts that the discretionary bonus components are incrementally informative about future financial performance but that the formula-based components are not predictive of future success after controlling for the current period's performance.

I follow the basic approach suggested by Hayes and Schaefer (2000) of regressing future financial performance on current compensation and current financial performance in a first-difference form. In particular, I estimate the following model:[19]

[19] The inferences are consistent when future stock return is considered as the dependent variable.

TABLE 3
Descriptive Statistics

Panel A: Discretionary Subsample

Variable[a]	n	Mean	Std	Q1	Median	Q3
Measures of Bonus Payment Components						
ΔTOTALBONUS	173	0.40	0.78	0.00	0.23	0.67
ΔFORMULABONUS	173	0.03	0.59	−0.20	0.00	0.29
Log(DISCBONUS)	173	−1.60	1.08	−2.27	−1.61	−0.88
Measures of Financial Performance						
ΔROA_{t+1}	211	0.012	0.086	−0.021	0.006	0.033
ΔROA_t	208	0.011	0.094	−0.016	0.008	0.036
ΔRET_t	169	−0.16	0.63	−0.44	−0.21	0.15
$\Delta Log(SALES_t)$	221	0.14	0.24	0.03	0.12	0.22
Measures of Contractible Performance						
FBONUS	168	0.84	0.74	0.00	0.83	1.30
ΔEPS	195	0.02	0.09	−0.01	0.01	0.05
Measure of Manipulability of Earnings						
SDTA	209	−3.18	1.03	−3.97	−3.19	−2.63
Measure of Executive Power						
CPS	229	0.41	0.14	0.31	0.41	0.48

Panel B: Non-Discretionary Subsample

Variable[a]	n	Mean	Std	Q1	Median	Q3
Measures of Contractible Performance						
FBONUS	349	1.06	0.70	0.65	1.00	1.41
ΔEPS	302	0.01	0.01	−0.01	0.01	0.03

(continued on next page)

Panel B: Non-Discretionary Subsample

Variable[a]	n	Mean	Std	Q1	Median	Q3
Measure of Manipulability of Earnings						
SDTA	304	−3.23	1.19	−3.95	−3.43	−2.78
Measures of Executive Power						
CPS	341	0.40	0.12	0.33	0.39	0.45

[a] See Table 2 for variable definitions.

$$\Delta ROA_{t+1} = \beta_0 + \beta_1(\Delta\text{Compensation variable}_t) + \beta_2 \Delta ROA_t + \beta_3 \Delta RET_t + \beta_3 \Delta Log(Sales_t) + \varepsilon \quad \text{(i)}$$

In terms of the independent variables, I focus on the financial performance in year t in order to economize on the number of parameters that have to be estimated.

Table 4 presents the results for the regression analyses.[20] I first estimate the regression using the change in total bonus payment as the measure of current compensation (model I). The coef-

TABLE 4

Future Financial Performance

	Dependent Variable[a]	
	ΔROA_{t+1}	
Independent Variables[a]	**I**	**II**
Intercept	−0.01	0.02
	(−0.57)	(1.08)
$\Delta TOTALBONUS_t$	0.01	
	(1.03)	
$\Delta FORMULABONUS_t$		0.00
		(0.29)
$Log(DISCBONUS_t)$		0.01**
		(2.49)
ΔROA_t	−0.24*	−0.25*
	(−1.68)	(−1.92)
ΔRET_t	0.01	0.01
	(0.77)	(1.05)
$\Delta\, Log(SALES_t)$	−0.07	−0.08
	(−1.25)	(−1.56)
Number of observations	98	98
Adj. R^2	16.76%	21.21%

*, **, *** Significant at 10 percent, 5 percent, and 1 percent, respectively, all based on two-tailed tests.

Reported are the coefficients from the models using heteroscedastic-consistent standard errors; t-statistics are in parentheses.

[a] $\Delta TOTALBONUS$ is the change in the salary-scaled total bonus payment. $\Delta FORMULABONUS$ is the change in the salary-scaled formula-based portion of the bonus payment. $Log(DISCBONUS_t)$ is the log of the salary-scaled discretionary component of the bonus payment. ΔROA is the change in return on assets, ΔRET is the change in market returns, and $\Delta log(SALES)$ is the change in the log of sales. In order to reduce the effect of extreme observations, I drop observations that fall in the 1st and 99th percentile of the distribution of the dependent variable. In addition to the variables listed, I include industry fixed effects based on one-digit SIC codes.

The variables are defined in Table 2.

[20] Some companies disclose in their From 8-K or Proxy Statement that the discretionary bonus was paid in relation with the achievement of a "special occasion" such as an IPO, a sale of company assets, a merger or an acquisition, a change in position, or the refinancing of debt. Table 4 is estimated using the sample that excludes such "special occasions." The results are qualitatively unchanged if the models in Table 4 are estimated on the full sample.

ficient on (*ΔTOTALBONUS*) is positive but not statistically significant.[21] As expected, the coefficient on the change in ROA is significantly negative.

In the second step, I decompose the measure (*ΔTOTALBONUS*) into the two components (*ΔFORMULABONUS*) and (*Log*(*DISCBONUS*)):[22] (*ΔTOTALBONUS*) is defined as [(formula bonus$_t$ + discretionary bonus$_t$) / salary$_t$] − [formula bonus$_{t-1}$ / salary$_{t-1}$], which can be separated into [formula bonus$_t$ / salary$_t$] − [formula bonus$_{t-1}$ / salary$_{t-1}$] (=*ΔFORMULABONUS*) and [(discretionary bonus$_t$) / salary$_t$] (=*DISCBONUS*). When the regression is estimated by including the two components of the change in total bonus in the regression separately, the coefficient on the discretionary portion is positive and significant, whereas the coefficient on the formula-based portion is insignificant.[23] These findings suggest that discretionary bonuses are incrementally predictive of future financial performance but that formula-based bonuses are not. In a broader sense, the results in Table 4 provide support for the notion that discretionary bonuses are based on non-contractible performance indicators.

In comparing Table 4 to the results in Hayes and Schaefer (2000), I view the analysis in this study as identifying one of the components of total cash compensation that drive the results in Hayes and Schaefer (2000).

Determinants of Discretionary Bonus Payments

Hypotheses 2 through 4 concern the occurrence of discretionary bonus payments. To investigate the predictions, I run a logit model using the indicator variable *DISCIND* as the dependent variable. Hypothesis 2 predicts that discretionary bonuses are more likely to occur when the outcome of the contractible measure falls in the tails of the distribution than when the contractible outcome is in the medium range. I use indicator variables in order to capture whether a company's contractible performance has an extreme value. In particular, I first employ an indicator variable, *TAILS*, which is set equal to 1 if the contractible outcome falls in the tails of the distribution. Specifically, when the contractible performance is measured by *FBONUS* (=(formula bonus/target bonus)), *TAILS* is set equal to 1 if *FBONUS* is 0 or if *FBONUS* falls in the top 15 percent of the distribution.[24] When the contractible performance is measured by *ΔEPS* (=(($EPS_t - EPS_{t-1}$) / Share Price$_{t-1}$)), *TAILS* is set equal to 1 if *ΔEPS* falls in the bottom or top 15 percent of the distribution.[25] In the second step, I estimate the logit model by splitting the indicator variable *TAILS* into its two components, *LOW* and *HIGH*, which capture whether the contractible outcome falls into the bottom or top end, respectively.

Hypotheses 3 and 4 are tested by including the variables capturing the manipulability of earnings (*SDTA*), and executive power (*CPS*) in the logit analyses. Specifically, I estimate the following models:

$$DISCIND = \beta_0 + \beta_1 TAILS + \beta_2 SDTA + \beta_3 CPS + \varepsilon; \quad \text{(ii)}$$

---

[21] Hayes and Schaefer (2000) run a related regression and obtain a significantly positive coefficient for the change in cash compensation. This discrepancy may be attributable to the much larger sample that is used in Hayes and Schaefer (2000).

[22] Since the distribution of *DISCBONUS* is highly skewed, I employ a log-transformation. The results are qualitatively unchanged if *DISCBONUS* is used without the log-transformation. Moreover, *DISCBONUS* is not differenced since there are no firms in the sample that disclosed payment of discretionary bonuses in two adjacent years.

[23] Untabulated results indicate that when the sample used in Table 4 is split into three groups based on the company's contractible performance as measured by (*ΔEPS*), the discretionary bonus payment is significantly more predictive of future financial performance when the company's contractible performance falls in the bottom third and when the company's contractible performance falls in the top third.

[24] I have chosen 15 percent since the number of observations for which *FBONUS* is zero comprise 15 percent of the sample. In other words, the number of observations in the two tails is chosen to be equal.

[25] 15 percent is chosen in order to facilitate comparison to the results using *FBONUS*.

$$DISCIND = \beta_0 + \beta_{1L}LOW + \beta_{1H}HIGH + \beta_2 SDTA + \beta_3 CPS + \varepsilon. \quad \text{(iii)}$$

Panel A of Table 5 shows the results of the logit regression analysis when the outcome of the contractible performance metric is measured by *FBONUS* (=formula bonus/target bonus), and Panel B provides the findings when the contractible metric is measured by ΔEPS ($=(EPS_t - EPS_{t-1})$ / Share Price$_{t-1}$).[26] Models Ia and Ib are estimated using the *TAILS* variable and models IIa and IIb are estimated using the two components of *TAILS, LOW*, and *HIGH*.

For the respective specifications Ia and Ib, the coefficients on *TAILS* are positive and statistically significant, indicating that companies are more likely to pay discretionary bonuses when the outcome of the contractible performance measure falls in the tails of the distribution. The odds ratios indicate that the probability of paying a discretionary bonus rises by approximately 4 (2) when the company's contractible performance falls in the tails of the distribution as measured by *FBONUS* (ΔEPS). Evaluated at mean values for *SDTA* and *CPS* (in model Ib), the probability of paying a discretionary bonus increases from approximately 18 percent to 50 percent (25 percent to 40 percent) when the company's contractible performance falls in the tails of the distribution as measured by *FBONUS* (ΔEPS). These findings provide some initial support for H2.

The results for models IIa and IIb indicate that companies are more likely to pay discretionary bonuses when their contractible performance is in either the bottom or the top tail of the distribution. Specifically, the coefficients on *LOW* and *HIGH* are positive and statistically significant in all specifications. The results in Panel A that are obtained when using *FBONUS* as the measure of contractible performance indicate that the odds ratio of paying a discretionary bonus rises by 6–7 when the company's contractible performance is in the lower tail. When the contractible performance is in the upper tail, the odds ratio of paying a discretionary bonus increases by a factor of approximately 2. Evaluated at mean values for *SDTA* and *CPS* (in model IIb), the probability of paying a discretionary bonus increases from approximately 18 percent to 60 percent (35 percent) when the company's contractible performance falls in the bottom (top) tail of the distribution. As can be observed, companies are more likely to pay discretionary bonuses when their contractible performance is low than when their contractible performance is high. The model analyzed above does not predict that discretionary bonuses are equally likely to occur for the low and for the high-contractible outcome. Specifically, the results of the simulation analysis indicate that, for the parameter ranges considered, it is more likely to observe discretionary bonus payments for the low-contractible outcome than for the high-contractible outcome. Thus, the discrepancy in the findings with respect to *LOW* and *HIGH* when *FBONUS* is used to capture the outcome of the contractible performance measure is consistent with the model that is analyzed above.

The results in Panel B that are obtained when using ΔEPS differ from the results in Panel A in that they indicate that there are no significant differences between the bottom and top tail. Specifically, the odds ratio of paying a discretionary bonus rises by approximately 2 when the company's contractible performance is in either the lower or upper tail. Evaluated at mean values for *SDTA* and *CPS* (in model IIb), the probability of paying a discretionary bonus increases from approximately 25 percent to 40 percent when the company's contractible performance is either very low or very high.

Overall, I interpret the findings in Table 5 as providing support for the premise that discretionary bonuses are more likely to occur when the outcome of the contractible performance measure falls in the tails of the distribution.

[26] Analogous to Table 4, Table 5 is estimated using the sample that excludes "special occasions." The results are qualitatively unchanged if the models in Table 5 are estimated on the full sample.

TABLE 5

Logit Analysis[a]

Panel A: Contractible Performance is Measured by *FBONUS* = (Formula bonus/target bonus)

Independent Variables[b]		Ia Coefficient/ [Odds Ratio] (t-statistic)	Ib Coefficient/ [Odds Ratio] (t-statistic)	IIa Coefficient/ [Odds Ratio] (t-statistic)	IIb Coefficient/ [Odds Ratio] (t-statistic)
Intercept	?	−1.97*** (−5.87)	−1.70** (−2.27)	−2.07*** (−6.07)	−1.87** (−2.49)
TAILS	+	1.40/[4.05]*** (6.05)	1.31/[3.69]*** (4.89)		
LOW	+			1.98/[7.27]*** (6.48)	1.84/[6.27]*** (5.30)
HIGH	+			0.75/[2.10]** (2.45)	0.72/[2.05]** (2.06)
SDTA	+		−0.00/[1.00] (−0.00)		−0.03/[0.97] (−0.29)
CPS	+		0.70/[2.01] (0.68)		0.64/[1.89] (0.63)
Number of observations		481	365	481	365
DISCBONUS > 0		132	104	132	104
Pseudo R^2		11.78%	10.81%	13.72%	12.35%

Panel B: Contractible Performance is Measured by $\Delta EPS = ((EPS_t - EPS_{t-1}) / Share\ Price_{t-1})$

Independent Variables[c]		Ia Coefficient/ [Odds Ratio] (t-statistic)	Ib Coefficient/ [Odds Ratio] (t-statistic)	IIa Coefficient/ [Odds Ratio] (t-statistic)	IIb Coefficient/ [Odds Ratio] (t-statistic)
Intercept	?	−1.42*** (−2.79)	−1.51** (−2.00)	−1.41*** (−2.79)	−1.49* (−1.96)

(continued on next page)

Panel B: Contractible Performance is Measured by $\Delta EPS = ((EPS_t - EPS_{t-1}) / Share\ Price_{t-1})$

Independent Variables[c]		Ia Coefficient/ [Odds Ratio] (t-statistic)	Ib Coefficient/ [Odds Ratio] (t-statistic)	IIa Coefficient/ [Odds Ratio] (t-statistic)	IIb Coefficient/ [Odds Ratio] (t-statistic)
TAILS	+	0.65/[1.91]*** (2.98)	0.69/[1.99]*** (2.81)		
LOW	+			0.66/[1.94]*** (2.60)	0.73/[2.07]** (2.54)
HIGH	+			0.63/[1.87]** (2.31)	0.64/[1.90]** (2.10)
SDTA	+		−0.11/[0.90] (−0.90)		−0.10/[0.90] (−0.86)
CPS	+		−0.47/[0.63] (−0.50)		−0.48/[0.62] (−0.52)
Number of observations		423	355	423	355
DISCBONUS > 0		138	116	138	116
Pseudo R^2		3.96%	4.86%	3.96%	4.87%

*, **, *** Significant at 10 percent, 5 percent, and 1 percent, respectively, all based on two-tailed tests.

Reported are the coefficients from the models with Z-statistics in parentheses. The models are estimated using Huber-White robust standard errors.

[a] The dependent variable for all models is *DISCIND*, which is equal to 1 if a discretionary bonus was paid, and 0 otherwise.

[b] *TAILS* is an indicator variable capturing whether the contractible performance has an extreme value. *TAILS* is set equal to 1 if *FBONUS* (= (Formula bonus/target bonus)) = 0 or if *FBONUS* falls in the top 15 percent of the distribution. *LOW* is set equal to 1 if *FBONUS* = 0. *HIGH* is set equal to 1 if *FBONUS* falls in the top 15 percent of the distribution. *SDTA* is the log of the standard deviation of the total accruals over the prior five years. *CPS* is the proportion of the total compensation of the top five executives that is captured by the CEO. In addition to the variables listed, I include industry fixed effects based on one-digit SIC codes.

[c] *TAILS* is an indicator variable capturing whether the contractible performance has an extreme value. *TAILS* is set equal to 1 if $\Delta EPS(=((EPS_t - EPS_{t-1}) / \text{Share Price}_{t-1}))$ falls in the top or bottom 15 percent of the distribution. *LOW* is set equal to 1 if ΔEPS falls in the bottom 15 percent of the distribution. *HIGH* is set equal to 1 if ΔEPS falls in the top 15 percent of the distribution. *SDTA* is the log of the standard deviation of the total accruals over the prior five years. *CPS* is the proportion of the total compensation of the top five executives that is captured by the CEO. In addition to the variables listed, I include industry fixed effects based on one-digit SIC codes.

The coefficients on the variable capturing accrual volatility (*SDTA*) are insignificant across all model specifications. Thus, the empirical setting used in the study does not provide support for the prediction based on the implicit contracting model in Baker et al. (1994) that discretionary bonuses are used more frequently when the contractible performance measures are more susceptible to manipulation.

Similarly, the coefficients on the variable capturing executive power (*CPS*) are insignificant across all model specifications, failing to support the hypothesis that discretionary bonuses occur more frequently in settings where management has considerable power. It is possible that the variable that is employed here is too imprecise a measure to detect a significant relationship.

Overall, the results from the logit analysis provide support for the hypothesis that discretionary bonuses are used more frequently when the outcome of the contractible performance measure falls in the tails of the distribution than when it falls in the medium range. The hypotheses that the occurrence of discretionary bonuses is positively related to the manipulability of the company's earnings or to the power of the executives are not supported.[27]

V. CONCLUSION

This study examines discretionary bonus payments. Building on the hypothesis that discretionary bonuses are paid based on non-contractible performance measures, I develop an analytical model that forms the basis for the primary hypothesis. The model, which builds on Rajan and Reichelstein (2009), yields the prediction that discretionary bonus payments occur when the outcome of the contractible measure is either low or high, but not when the outcome of the contractible measure falls in the medium range.

I also test whether discretionary bonus payments are related to the manipulability of the contractible measures—a prediction that follows from the model in Baker et al. (1994). I contrast the predictions from agency theory against a prediction based on managerial power theory.

The empirical results provide evidence that discretionary bonus payments are related to non-contractible measures that are related to future financial performance. Moreover, the analyses yield support for the prediction from the model that discretionary bonus payments occur significantly more often when the outcome of the contractible measure falls in the tails of the distribution. The findings do not provide support for the prediction that there is a positive relationship between discretionary bonus payments and the manipulability of the contractible performance measures. The findings in this study also do not provide support for the hypothesis that discretionary bonuses are positively related to executive power.

This study is subject to some limitations. The sample collected from public sources may be subject to selection bias. Although companies are required to disclose discretionary bonus payments in a Form 8-K or a proxy statement, the regulation was still fairly new during the sample period, such that it is possible that some companies did not disclose the information. Second, the sample size is relatively small. It is conceivable that the lack of findings with respect to the manipulability of earnings and executive power is, at least partially, attributable to the low statistical power generated by the small sample. Moreover, it is possible that there are determinants of discretionary bonus payments that are not considered in this study. Such omitted variables would raise concerns with respect to the empirical findings of the study.

[27] The literature has documented that bonus contracts of loss-making firms are structured differently than bonus contracts of profitable firms (Matejka et al. 2009). The results are robust to excluding loss-making firms from the sample.

APPENDIX A

Proof of Proposition (1)

I proceed by characterizing the solution to program P1. Let λ and μ denote the Lagrange-multipliers for constraints (i) and (ii), respectively. Denote by β^{jk} the Lagrange-multiplier for the constraint $w^j - s^{jk} \geq 0$ (there are four such inequalities embedded in constraint (iii)). The first-order conditions for the optimal solution with regard to w^h and w^l are:

$$-p_h^h + \beta^{hh} + \beta^{hl} = 0; \tag{2}$$

$$-(1 - p_h^h) + \beta^{lh} + \beta^{ll} = 0. \tag{3}$$

The first-order conditions for the optimal solution with regard to s^{hh}, s^{hl}, s^{lh}, and s^{ll} are:

$$\lambda \cdot p_h^h \cdot q_{hh}^h \cdot U'(s^{hh}) + \mu \cdot [p_h^h \cdot q_{hh}^h - p_h^l \cdot q_{hh}^l] \cdot U'(s^{hh}) - \beta^{hh} = 0; \tag{4}$$

$$\lambda \cdot p_h^h \cdot (1 - q_{hh}^h) \cdot U'(s^{hl}) + \mu \cdot [p_h^h \cdot (1 - q_{hh}^h) - p_h^l \cdot (1 - q_{hh}^l)] \cdot U'(s^{hl}) - \beta^{hl} = 0; \tag{5}$$

$$\lambda \cdot (1 - p_h^h) \cdot q_{lh}^h \cdot U'(s^{lh}) + \mu \cdot [(1 - p_h^h) \cdot q_{lh}^h - (1 - p_h^l) \cdot q_{lh}^l] \cdot U'(s^{lh}) - \beta^{lh} = 0; \tag{6}$$

$$\lambda \cdot (1 - p_h^h) \cdot (1 - q_{lh}^h) \cdot U'(s^{ll}) + \mu \cdot [(1 - p_h^h) \cdot (1 - q_{lh}^h) - (1 - p_h^l) \cdot (1 - q_{lh}^l)] \cdot U'(s^{ll}) - \beta^{ll} = 0. \tag{7}$$

By (4) and (5), and MLRP:

$$\frac{\beta^{hh}}{p_h^h \cdot q_{hh}^h \cdot U'(s^{hh})} = \lambda + \mu \cdot \left[1 - \frac{p_h^l \cdot q_{hh}^l}{p_h^h \cdot q_{hh}^h}\right] > \lambda + \mu \cdot \left[1 - \frac{p_h^l \cdot (1 - q_{hh}^l)}{p_h^h \cdot (1 - q_{hh}^h)}\right] = \frac{\beta^{hl}}{p_h^h \cdot (1 - q_{hh}^h) \cdot U'(s^{hl})}. \tag{8}$$

By (6) and (7), and MLRP:

$$\frac{\beta^{lh}}{(1 - p_h^h) \cdot q_{lh}^h \cdot U'(s^{lh})} = \lambda + \mu \cdot \left[1 - \frac{(1 - p_h^l) \cdot q_{lh}^l}{(1 - p_h^h) \cdot q_{lh}^h}\right] > \lambda + \mu \cdot \left[1 - \frac{(1 - p_h^l) \cdot (1 - q_{lh}^l)}{(1 - p_h^h) \cdot (1 - q_{lh}^h)}\right] = \frac{\beta^{ll}}{(1 - p_h^h) \cdot (1 - q_{lh}^h) \cdot U'(s^{ll})}. \tag{9}$$

It is clear that $w^h = \max\{s^{hh}, s^{hl}\}$ and $w^l = \max\{s^{lh}, s^{ll}\}$. To show that $s^{hh} \geq s^{hl}$, suppose that the opposite were true. Then, $s^{hh} < s^{hl} = w^h \Rightarrow \beta^{hh} = 0$ (by complementary slackness), which in turn would imply $\beta^{hl} = p^h$ (from (2)). But from (8), $\beta^{hh} = 0 \Rightarrow \beta^{hl} < 0$, which is a contradiction. We must therefore have $w^h = s^{hh} \geq s^{hl}$. To show that $s^{lh} \geq s^{ll}$, suppose that the opposite were true. Then, $s^{lh} < s^{ll} = w^l \Rightarrow \beta^{lh} = 0$ (by complementary slackness), which in turn would imply $\beta^{ll} = (1 - p_h^h)$ (from (3)). But from (9), $\beta^{lh} = 0 \Rightarrow \beta^{ll} < 0$, which is a contradiction. We must therefore have $w^l = s^{lh} \geq s^{ll}$. Using these results, the principal's problem P1 can be re-cast as follows:

$$\underset{w^h, w^l, \delta, \varepsilon}{\text{Min}} \; [w^h \cdot p_h^h + w^l \cdot (1 - p_h^h)]$$

subject to:

$$[U(w^h) \cdot q_{hh}^h + (U(w^h) - \varepsilon) \cdot (1 - q_{hh}^h)] \cdot p_h^h + [U(w^l) \cdot q_{lh}^h + (U(w^l) - \delta) \cdot (1 - q_{lh}^h)] \cdot (1 - p_h^h) - e(a^h) - \bar{U} \geq 0, \tag{10}$$

$$[U(w^h)\cdot q^h_{hh}+(U(w^h)-\varepsilon)\cdot(1-q^h_{hh})]\cdot p^h_h+[U(w^l)\cdot q^h_{lh}+(U(w^l)-\delta)\cdot(1-q^h_{lh})]\cdot(1-p^h_h)$$
$$-e(a^h)-\{[U(w^h)\cdot q^l_{hh}+(U(w^h)-\varepsilon)\cdot(1-q^l_{hh})]\cdot p^l_h+[U(w^l)\cdot q^l_{lh}+(U(w^l)-\delta)\cdot(1$$
$$-q^l_{lh})]\cdot(1-p^l_h)-e(a^l)\}\geq 0, \tag{11}$$

$$\delta\geq 0, \tag{12}$$

$$\varepsilon\geq 0, \tag{13}$$

where (10) is the agent's (IR) constraint and (11) is the agent's (IC) constraint. Since both (10) and (11) are satisfied as equalities, it must be the case that:

$$u^h=U(w^h)=\bar{U}+\frac{e(a^h)\cdot(1-p^l_h)-e(a^l)\cdot(1-p^h_h)}{(p^h_h-p^l_h)}-\delta\cdot\frac{(1-p^h_h)\cdot(1-p^l_h)\cdot(q^h_{lh}-q^l_{lh})}{(p^h_h-p^l_h)}$$
$$-\varepsilon\cdot\frac{p^l_h\cdot(1-p^h_h)\cdot(1-q^l_{hh})-p^h_h\cdot(1-p^l_h)\cdot(1-q^h_{hh})}{(p^h_h-p^l_h)} \tag{14}$$

and:

$$u^l=U(w^l)=\bar{U}+\frac{e(a^l)\cdot p^h_h-e(a^h)\cdot p^l_h}{(p^h_h-p^l_h)}+\delta\cdot\frac{p^h_h\cdot(1-p^l_h)\cdot(1-q^l_{lh})-p^l_h\cdot(1-p^h_h)\cdot(1-q^h_{lh})}{(p^h_h-p^l_h)}$$
$$+\varepsilon\cdot\frac{p^l_h\cdot p^h_h\cdot(q^h_{hh}-q^l_{hh})}{(p^h_h-p^l_h)}. \tag{15}$$

Thus, the problem in program P1 can be reduced to the following program:

Program P3:

$$\underset{\delta,\varepsilon}{\text{Min}}[Z(\delta,\varepsilon)]=\underset{\delta,\varepsilon}{\text{Min}}[p^h_h\cdot U^{-1}(u^h)+(1-p^h_h)\cdot U^{-1}(u^l)]$$

subject to:
(A1) $\delta\geq 0$,
(A2) $\varepsilon\geq 0$,
where $U^{-1}(\cdot)$ is the inverse of the agent's utility function, and u^h and u^l are given by the expressions in (14) and (15).

To show that the problem is well defined, we can show that Z is convex in (δ,ε):

$$\frac{\partial^2 Z(\delta,\varepsilon)}{\partial\delta^2}=p^h_h\cdot V'(u^h)\cdot\left[\frac{(1-p^h_h)\cdot(1-p^l_h)\cdot(q^h_{lh}-q^l_{lh})}{(p^h_h-p^l_h)}\right]^2$$
$$+(1-p^h_h)\cdot V'(u^l)\cdot\left[\frac{p^h_h\cdot(1-p^l_h)\cdot(1-q^l_{lh})-p^l_h\cdot(1-p^h_h)\cdot(1-q^h_{lh})}{(p^h_h-p^l_h)}\right]^2$$

$$\frac{\partial^2 Z(\delta,\varepsilon)}{\partial\varepsilon^2}=p^h_h\cdot V'(u^h)\cdot\left[\frac{p^l_h\cdot(1-p^h_h)\cdot(1-q^l_{hh})-p^h_h\cdot(1-p^l_h)\cdot(1-q^h_{hh})}{(p^h_h-p^l_h)}\right]^2$$
$$+(1-p^h_h)\cdot V'(u^l)\cdot\left[\frac{p^h_h\cdot p^l_h\cdot(q^h_{hh}-q^l_{hh})}{(p^h_h-p^l_h)}\right]^2$$

$$\frac{\partial^2 Z(\delta,\varepsilon)}{\partial\delta\partial\varepsilon} = p_h^h \cdot V'(u^h) \cdot \left[\frac{(1-p_h^h)\cdot(1-p_h^l)\cdot(q_{lh}^h - q_{lh}^l)}{(p_h^h - p_h^l)}\right] \cdot \left[\frac{p_h^l\cdot(1-p_h^h)\cdot(1-q_{hh}^l) - p_h^h\cdot(1-p_h^l)\cdot(1-q_{hh}^h)}{(p_h^h - p_h^l)}\right] + (1-p_h^h)\cdot V'(u^l)\cdot\left[\frac{p_h^h\cdot p_h^l\cdot(q_{hh}^h - q_{hh}^l)}{(p_h^h - p_h^l)}\right] \cdot \left[\frac{p_h^h\cdot(1-p_h^l)\cdot(1-q_{lh}^l) - p_h^l\cdot(1-p_h^h)\cdot(1-q_{lh}^h)}{(p_h^h - p_h^l)}\right]$$

$$\frac{\partial^2 Z(\delta,\varepsilon)}{\partial\delta^2}\cdot\frac{\partial^2 Z(\delta,\varepsilon)}{\partial\varepsilon^2} - \left[\frac{\partial^2 Z(\delta,\varepsilon)}{\partial\delta\partial\varepsilon}\right]^2 = -\frac{(-1+p_h^h)\cdot p_h^h\cdot(p_h^l\cdot(-1+q_{lh}^h)+p_h^h\cdot((-1+q_{hh}^h)\cdot(-1+q_{lh}^l)+p_h^l}{(p_h^h - p_h^l)^2} \cdot \frac{(-q_{lh}^h + q_{hh}^h + q_{lh}^l - q_{hh}^h\cdot q_{lh}^l)) + (-1+p_h^h)\cdot p_h^l\cdot(-1+q_{lh}^h)\cdot\ q_{hh}^l)^2\cdot V'(u^h)\cdot V'(u^l)}{(p_h^h - p_h^l)^2} \geq 0.$$

The first-order conditions for the optimal solution with regard to δ and ε are:

$$-\frac{\partial Z(\delta,\varepsilon)}{\partial\delta} + \beta_1 = 0, \tag{16}$$

$$-\frac{\partial Z(\delta,\varepsilon)}{\partial\varepsilon} + \beta_2 = 0, \tag{17}$$

where:

$$\frac{\partial Z(\delta,\varepsilon)}{\partial\delta} = p_h^h\cdot V(u^h)\cdot(-1)\cdot\left[\frac{(1-p_h^h)\cdot(1-p_h^l)\cdot(q_{lh}^h - q_{lh}^l)}{(p_h^h - p_h^l)}\right] + (1-p_h^h)\cdot V(u^l)\cdot\left[\frac{p_h^h\cdot(1-p_h^l)\cdot(1-q_{lh}^l) - p_h^l\cdot(1-p_h^h)\cdot(1-q_{lh}^h)}{(p_h^h - p_h^l)}\right], \tag{18}$$

$$\frac{\partial Z(\delta,\varepsilon)}{\partial\varepsilon} = p_h^h\cdot V(u^h)\cdot(-1)\cdot\left[\frac{p_h^l\cdot(1-p_h^h)\cdot(1-q_{hh}^l) - p_h^h\cdot(1-p_h^l)\cdot(1-q_{hh}^h)}{(p_h^h - p_h^l)}\right] + (1-p_h^h)\cdot V(u^l)\cdot\left[\frac{p_h^h\cdot p_h^l\cdot(q_{hh}^h - q_{hh}^l)}{(p_h^h - p_h^l)}\right]. \tag{19}$$

$V(\cdot)$ represents the derivative of $U^{-1}(\cdot)$ and β_1, β_2 are the Lagrange-multipliers for constraint (A1) and (A2), respectively.

$\beta_1 \geq 0$ and $\beta_2 \geq 0$ implies $\frac{\partial Z(\delta,\varepsilon)}{\partial\delta} \geq 0$ and $\frac{\partial Z(\delta,\varepsilon)}{\partial\varepsilon} \geq 0$ (by (16) and (17)), which in turn implies the following (by (18) and (19)):

$$\frac{V(u^l)}{V(u^h)} \geq \frac{p_h^h\cdot(1-p_h^l)\cdot(q_{lh}^h - q_{lh}^l)}{p_h^h\cdot(1-p_h^l)\cdot(1-q_{lh}^l) - p_h^l\cdot(1-p_h^h)\cdot(1-q_{lh}^h)}; \tag{20}$$

$$\frac{V(u^l)}{V(u^h)} \geq \frac{p^l \cdot (1-p^h) \cdot (1-q_h^l) - p^h \cdot (1-p^l) \cdot (1-q_h^h)}{(1-p^h) \cdot p^l \cdot (q_h^h - q_h^l)}. \tag{21}$$

It suffices to show that under the conditions in (1), $\varepsilon^* > 0$.

(21) and (1) yield the following relationship:

$$\frac{V(u^l)}{V(u^h)} \geq \frac{p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}{(1-p_h^h) \cdot p_h^l \cdot (q_{hh}^h - q_{hh}^l)} > \frac{V(l)}{V(h)},$$

which implies:

$$\frac{V(u^l)}{V(u^h)} > \frac{V(l)}{V(h)}, \tag{22}$$

where:

$$V(l) = V\left(\bar{U} + \frac{e(a^l) \cdot p_h^h - e(a^h) \cdot p_h^l}{(p_h^h - p_h^l)}\right), \text{ and}$$

$$V(h) = V\left(\bar{U} + \frac{e(a^h) \cdot (1-p_h^l) - e(a^l) \cdot (1-p_h^h)}{(p_h^h - p_h^l)}\right).$$

To show that $\varepsilon^* > 0$, suppose that $\varepsilon^* = 0$. Then $\frac{V(u^l)}{V(u^h)} = \frac{V(l)}{V(h)}$, violating (22). Therefore, $\varepsilon^* \neq 0$.

It follows that $\beta_2 = 0$ and therefore, $\frac{\partial Z(\delta,\varepsilon)}{\partial \varepsilon}\Big|_{\delta=0} = 0$:

$$\frac{\partial Z(\delta,\varepsilon)}{\partial \varepsilon}\bigg|_{\delta=0} = p_h^h \cdot V(H') \cdot (-1) \cdot \left[\frac{p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}{(p_h^h - p_h^l)}\right]$$
$$+ (1-p_h^h) \cdot V(L') \cdot \left[\frac{p_h^h \cdot p_h^l \cdot (q_{hh}^h - q_{hh}^l)}{(p_h^h - p_h^l)}\right] = 0$$

or, equivalently:

$$\frac{V(L')}{V(H')} = \frac{p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}{(1-p_h^h) \cdot p_h^l \cdot (q_{hh}^h - q_{hh}^l)},$$

where:

$$V(L') = V\left(\bar{U} + \frac{e(a^l) \cdot p_h^h - e(a^h) \cdot p_h^l}{(p_h^h - p_h^l)} + \varepsilon \cdot \frac{p_h^l \cdot p_h^h \cdot (q_{hh}^h - q_{hh}^l)}{(p_h^h - p_h^l)}\right), \text{ and}$$

$$V(H') = V\left(\bar{U} + \frac{e(a^h) \cdot (1-p_h^l) - e(a^l) \cdot (1-p_h^h)}{(p_h^h - p_h^l)}\right.$$
$$\left. - \varepsilon \cdot \frac{p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}{(p_h^h - p_h^l)}\right).$$

In order to show that there exists a $\varepsilon^* > 0$, it suffices to show that (i) $R(\varepsilon) = \frac{V(L')}{V(H')}$ is increasing in ε and that (ii) there exists a $\hat{\varepsilon} > 0$ such that:

$$R(\hat{\varepsilon}) \geq \frac{p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}{(1-p_h^h) \cdot p_h^l \cdot (q_{hh}^h - q_{hh}^l)}.$$

Part (i)

$R(\varepsilon) = \frac{V(L')}{V(H')}$ is increasing in ε:

$$\frac{\partial R(\varepsilon)}{\partial \varepsilon} = \frac{V'(L') \cdot \left(\frac{p_h^l \cdot p_h^h \cdot (q_{hh}^h - q_{hh}^l)}{(p_h^h - p_h^l)} \right) \cdot V(H')}{V(H')^2}$$

$$+ \frac{V(L') \cdot V'(H') \cdot \left(\frac{p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}{(p_h^h - p_h^l)} \right)}{V(H')^2} \geq 0.$$

(We know that $(p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h))$ is positive by (1)).

Part (ii)

Note that:

$$\frac{p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}{(1-p_h^h) \cdot p_h^l \cdot (q_{hh}^h - q_{hh}^l)} < 1.$$

Moreover, there exists a $\hat{\varepsilon} > 0$ such that $L' = H'$ and therefore $\frac{V(L')}{V(H')} = 1$:

$$\hat{\varepsilon} = \frac{e(a^h) - e(a^l)}{p_h^l \cdot p_h^h \cdot (q_{hh}^h - q_{hh}^l) + p_h^l \cdot (1-p_h^h) \cdot (1-q_{hh}^l) - p_h^h \cdot (1-p_h^l) \cdot (1-q_{hh}^h)}.$$

(L', H' are >0 when $\varepsilon = \hat{\varepsilon}$).

Therefore, there exists a (unique) $\varepsilon^* > 0$ such that $\frac{\partial Z(\delta, \varepsilon^*)}{\partial \varepsilon}\Big|_{\delta=0} = 0$. ■

Proof of Proposition (2)

It suffices to show that the solution to program P2 is such that $w^m = s^{mh} = s^{ml}$. Let λ and μ denote the Lagrange-multipliers for constraints (i) and (ii), respectively. Denote by β^{jk} the Lagrange-multiplier for the constraint $w^j - s^{jk} \geq 0$ (there are six such inequalities embedded in constraint (iii)). The first-order conditions for the optimal solution with regard to w^h, w^m, and w^l are:

$$-p_h^h + \beta^{hh} + \beta^{hl} = 0, \tag{23}$$

$$-p_m^h + \beta^{mh} + \beta^{ml} = 0, \tag{24}$$

$$-(1-p_h^h - p_m^h) + \beta^{lh} + \beta^{ll} = 0. \tag{25}$$

Similarly, the first-order conditions for the optimal solution with regard to s^{hh}, s^{hl}, s^{mh}, s^{ml}, s^{lh}, and s^{ll} are:

$$\lambda \cdot p_h^h \cdot q_{hh}^h \cdot U'(s^{hh}) + \mu \cdot [p_h^h \cdot q_{hh}^h - p_h^l \cdot q_{hh}^l] \cdot U'(s^{hh}) - \beta^{hh} = 0, \tag{26}$$

$$\lambda \cdot p_h^h \cdot (1-q_{hh}^h) \cdot U'(s^{hl}) + \mu \cdot [p_h^h \cdot (1-q_{hh}^h) - p_h^l \cdot (1-q_{hh}^l)] \cdot U'(s^{hl}) - \beta^{hl} = 0, \tag{27}$$

$$\lambda \cdot p_m^h \cdot q_{mh}^h \cdot U'(s^{mh}) + \mu \cdot [p_m^h \cdot q_{mh}^h - p_m^l \cdot q_{mh}^l] \cdot U'(s^{mh}) - \beta^{mh} = 0, \tag{28}$$

$$\lambda \cdot p_m^h \cdot (1 - q_{mh}^h) \cdot U'(s^{ml}) + \mu \cdot [p_m^h \cdot (1 - q_{mh}^h) - p_m^l \cdot (1 - q_{mh}^l)] \cdot U'(s^{ml}) - \beta^{ml} = 0, \tag{29}$$

$$\lambda \cdot (1 - p_h^h - p_m^h) \cdot q_{lh}^h \cdot U'(s^{lh}) + \mu \cdot [(1 - p_h^h - p_m^h) \cdot q_{lh}^h - (1 - p_h^l - p_m^l) \cdot q_{lh}^l] \cdot U'(s^{lh}) - \beta^{lh} = 0, \tag{30}$$

$$\lambda \cdot (1 - p_h^h - p_m^h) \cdot (1 - q_{lh}^h) \cdot U'(s^{ll}) + \mu \cdot [(1 - p_h^h - p_m^h) \cdot (1 - q_{lh}^h) - (1 - p_h^l - p_m^l) \cdot (1 - q_{lh}^l)] \cdot U'(s^{ll}) - \beta^{ll} = 0. \tag{31}$$

By (28) and (29), and MLRP:

$$\frac{\beta^{mh}}{p_m^h \cdot q_{lh}^h \cdot U'(s^{mh})} = \lambda + \mu \cdot \left[1 - \frac{p_m^l \cdot q_{mh}^l}{p_m^h \cdot q_{mh}^h}\right] > \lambda + \mu \cdot \left[1 - \frac{p_m^l \cdot (1 - q_{mh}^l)}{p_m^h \cdot (1 - q_{mh}^h)}\right] = \frac{\beta^{ml}}{p_m^h \cdot (1 - q_{lh}^h) \cdot U'(s^{ml})}. \tag{32}$$

By (29) and (31), MLRP, and assuming that $q_{lh}^h \geq q_{mh}^h$ and $q_{mh}^l \geq q_{lh}^l$:

$$\frac{\beta^{ml}}{p_m^h \cdot (1 - q_{mh}^h) \cdot U'(s^{ml})} = \lambda + \mu \cdot \left[1 - \frac{p_m^l \cdot (1 - q_{mh}^l)}{p_m^h \cdot (1 - q_{mh}^h)}\right] > \lambda + \mu \cdot \left[1 - \frac{p_l^l \cdot (1 - q_{lh}^l)}{p_l^h \cdot (1 - q_{lh}^h)}\right] = \frac{\beta^{ll}}{p_l^h \cdot (1 - q_{lh}^h) \cdot U'(s^{ll})}. \tag{33}$$

It is clear that $w^m = \max\{s^{mh}, s^{ml}\}$. To show that $s^{mh} \geq s^{ml}$, suppose that the opposite were true. Then, $s^{mh} < s^{ml} = w^m \Rightarrow \beta^{mh} = 0$ (by complementary slackness), which in turn would imply $\beta^{ml} = p_m^h$ (from (24)). But from (32), $\beta^{mh} = 0 \Rightarrow \beta^{ml} < 0$, which is a contradiction.

To show that $w^m = s^{mh} = s^{ml}$, suppose not, so that $w^m = s^{mh} > s^{ml}$. This implies $\beta^{ml} = 0$ (by complementary slackness). But from (33), $\beta^{ml} = 0 \Rightarrow \beta^{ll} < 0$, which is impossible. ■

APPENDIX B
SUMMARY OF THE MONTE CARLO SIMULATIONS

General Model Specifications

p_j^i, q_{jh}^i — In all models, the probabilities can be randomly chosen from (0, 1), subject to the MLRP constraints, the constraints that $q_{hh}^h \geq q_{mh}^h \geq q_{lh}^h$, $q_{hh}^l \geq q_{mh}^l \geq q_{lh}^l$, $(q_{hh}^h - q_{mh}^h) \geq (q_{mh}^h - q_{lh}^h)$, and $(q_{hh}^l - q_{mh}^l) \leq (q_{mh}^l - q_{lh}^l)$.

$e(a^l)$ — The cost of taking the low action is set to zero in all models.

$U(s) = \frac{1}{1-\gamma} s^{(1-\gamma)}$ — I assume a power utility function. The parameter γ varies across the models.

Model Parameters			Results			
γ	$e(a^h)$	$\bar{U}$	Number of Datapoints Generated	Number of Cases Where y is Used for x^l	Number of Cases Where y is Used for x^m	Number of Cases Where y is Used for x^h
0.5	$\in[1,10]$	$\in[5,35]$	716	272	0	71

(continued on next page)

Model Parameters			Results			
γ	$e(a^h)$	$\bar{U}$	Number of Datapoints Generated	Number of Cases Where y is Used for x^l	Number of Cases Where y is Used for x^m	Number of Cases Where y is Used for x^h
0.5	$\in[1,5]$	$\in[5,25]$	615	194	0	50
$\in[0.1,0.9]$	$\in[1,10]$	$\in[5,35]$	206	73	0	13

REFERENCES

Baiman, S., and M. Rajan. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review* 70 (October): 557–579.

———, D. Larcker, and M. Rajan. 1995. Organizational design for business units. *Journal of Accounting Research* 33 (Autumn): 205–229.

Baker, G., R. Gibbons, and K. Murphy. 1994. Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics* 109 (November): 1125–1156.

Bebchuk, L., M. Cremers, and U. Peyers. 2008. CEO centrality. Working paper, Harvard Law School.

———, J. Fried, and D. Walker. 2002. Managerial power and rent extraction in the design of executive compensation. NBER Working Paper. Cambridge, MA: NBER.

Boyd, B. 1994. Board control and CEO compensation. *Strategic Management Journal* 15 (June): 335–344.

Budde, J. 2007. Performance measure congruity and the balanced scorecard. *Journal of Accounting Research* 45 (June): 515–539.

Bull, C. 1987. The existence of self-enforcing implicit contracts. *The Quarterly Journal of Economics* 102 (February): 147–159.

Bushman, R., R. Indjejikian, and A. Smith. 1996. CEO compensation: The role of individual performance evaluation. *Journal of Accounting and Economics* 21 (April): 161–193.

Core, J., W. Guay, and R. Verrecchia. 2003. Price versus non-price performance measures in optimal CEO compensation contracts. *The Accounting Review* 78 (October): 957–981.

———, R. Holthausen, and D. Larcker. 1999. Corporate governance, chief executive officer compensation and firm performance. *Journal of Financial Economics* 51 (March): 371–406.

DeAngelo, L. 1986. Accounting numbers as market valuation substitutes: A study of management buyouts of public shareholders. *The Accounting Review* 61 (July): 400–420.

Dechow, P., and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 35–59.

Gibbs, M., K. Merchant, W. Van der Stede, and M. Vargus. 2004. Determinants and effects of subjectivity in incentives. *The Accounting Review* 79 (April): 409–436.

Hayes, R., and S. Schaefer. 2000. Implicit contracts and the explanatory power of top executive compensation for future performance. *The Rand Journal of Economics* 31 (Summer): 273–293.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (April): 85–107.

Indjejikian, R., and D. Nanda. 2002. Executive target bonuses and what they imply about performance standards. *The Accounting Review* 77 (October): 793–819.

———, and M. Matejka. 2006. Organizational slack in decentralized firms: the role of business unit controllers. *The Accounting Review* 81 (July): 849–872.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (Autumn): 193–228.

Lambert, R., D. Larcker, and K. Weigelt. 1993. The structure of organizational incentives. *Administrative Science Quarterly* 38 (September): 438–461.

Leone, A., and S. Rock. 2002. Empirical tests of budget ratcheting and its effect on managers' discretionary accrual choices. *Journal of Accounting and Economics* 33 (February): 43–67.

Levin, J. 2003. Relational incentive contracts. *The American Economic Review* 93 (June): 835–847.

MacLeod, W. 2003. Optimal contracting with subjective evaluation. *The American Economic Review* 93 (March): 216–240.

Matejka, M., K. Merchant, and W. Van der Stede. 2009. Employment horizon and the choice of performance measures: Empirical evidence from annual bonus plans of loss-making entities. *Management Science* 55 (June): 890–905.

Merchant, K., and J. Manzoni. 1989. The achievability of budget targets in profit centers: A field study. *The Accounting Review* 64 (July): 539–558.

Murphy, K. 2000. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30 (December): 245–278.

———, and P. Oyer. 2003. Discretion in executive incentive contracts. Working paper, University of Southern California and Stanford University.

Pearce, D., and E. Stacchetti. 1998. The interaction of implicit and explicit contracts in repeated agency. *Games and Economic Behavior* 23 (April): 75–96.

Rajan, M., and S. Reichelstein. 2006. Subjective performance indicators and discretionary bonus pools. *Journal of Accounting Research* 44 (June): 585–618.

———, and ———. 2009. Objective versus subjective indicators of managerial performance. *The Accounting Review* 84 (January): 209–237.

Sloan, R. 1993. Accounting earnings and top executive compensation. *Journal of Accounting and Economics* 16 (January): 55–100.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 1951–1984

American Accounting Association
DOI: 10.2308/accr.2010.85.6.1951

Once Bitten, Twice Shy: The Relation between Outcomes of Earnings Guidance and Management Guidance Strategy

Mei Feng
University of Pittsburgh
Adam S. Koch
University of Virginia

ABSTRACT: We examine how management quarterly guidance strategy is affected by various outcomes from previously issued guidance. We find that managers are less likely to provide quarterly earnings guidance for a given year when past management forecasts have been overly optimistic, when past forecasts were unsuccessful at influencing analysts' expectations, when past forecasts failed to reduce information asymmetry, and when past forecasts resulted in earnings disappointments. For firms that continue to give guidance, adverse prior outcomes also affect the precision of future guidance and the number of quarters within a year for which they give guidance.

Keywords: *earnings guidance; management forecasts; corporate disclosure.*

Data Availability: *Data used in this study are available from public sources.*

I. INTRODUCTION

This study examines how management quarterly guidance strategy is affected by outcomes from previously issued guidance. Forecasting earnings is fundamentally a multi-period process in which outcomes from one period influence managerial, investor, and analyst decisions about future forecasts. However, a recent review by Hirst et al. (2008) emphasizes that most prior research ignores the extent to which forecast outcomes from one period influence forecast antecedents and characteristics in subsequent periods. Our study begins to fill this void by documenting the extent to which a wide range of adverse outcomes affects future guidance be-

We appreciate the helpful comments of Beverly Walther (the editor), an anonymous reviewer, Yonca Ertimur, Harry Evans, Robert Freeman, Weili Ge, Guojin Gong, Vicky Hoffman, Craig Lefanowicz, Chan Li, Yinghua Li, Roger Martin, Sarah McVay, Don Moser, Nandu Nagarajan, Susan Porter, Doug Skinner, and workshop participants at American University, Carnegie Mellon University, Fordham University, George Mason University, The George Washington University, Southern Methodist University, University at Buffalo, University of New Hampshire, University of Pittsburgh, The University of Texas at Dallas, and University of Virginia. Professor Koch gratefully acknowledges financial support from the Ernst & Young Foundation and the McIntire School of Commerce.

Editor's note: Accepted by Beverly Walther.

Submitted: April 2008
Accepted: May 2010
Published Online: November 2010

havior. This is an important contribution because, as Hirst et al. (2008) argue, exploring the multi-period nature of forecasting is necessary to extend our understanding of forecasting beyond immediate one-time decisions and reactions.

We adopt a framework in which managers seek to maximize shareholder value by aligning market expectations with managers' private information.[1] Our framework divides progress toward attaining this goal into four outcomes. First, managers want to provide attainable guidance that can be met or beaten in order to minimize legal liability (Kasznik 1999) and develop a reputation for credible reporting (Graham et al. 2005). Second, guidance is only effective when market participants follow managers' guidance. Third, guidance is intended to reduce information asymmetry (Verrecchia 2001). Fourth, guidance is used by managers to avoid disappointing market participants at earnings announcements (Richardson et al. 2004; Cotter et al. 2006).

We argue that failing to achieve desired outcomes affects guidance strategy by influencing managers' expectations about the benefits and costs of future guidance. Benefits to providing earnings guidance include reducing litigation risk (e.g., Skinner 1994, 1997), building a reputation for transparent reporting (e.g., Graham et al. 2005; Hutton and Stocken 2009), reducing information asymmetry (Diamond and Verrecchia 1991), and guiding analysts' forecasts toward beatable earnings targets (e.g., Cotter et al. 2006), while costs to guidance include legal liability for optimistic forecasting (Kasznik 1999). Firms can respond to changes in the expected costs and benefits in a number of ways. For example, a firm that experienced adverse outcomes could conclude that forecasting is no longer in their interest, choosing not to forecast again. Other firms could choose to continue providing guidance, but reduce intra-year forecast frequency or change the form of guidance (e.g., providing a range of possible values, rather than a point estimate), improving *ex post* "accuracy" by reducing *ex ante* guidance precision. Alternatively, firms could choose not to change their forecasting strategy based on a short period of adverse outcomes. Given the costs and benefits attributed to guidance, it remains an open empirical question as to how adverse past outcomes influence managers' forecasting behavior.

We empirically investigate how adverse outcomes are associated with three dimensions of future quarterly forecasting strategy: whether to guide at all next period, the number of fiscal quarters for which firms give guidance, and the specificity of guidance provided. We find that falling short of one's own guidance, failing to align market expectations with management guidance, failing to reduce information asymmetry, or disappointing market participants at earnings announcements increases the probability of omitting guidance in the next year. For firms that continue to give guidance, adverse prior outcomes are also associated with the precision of future guidance and the number of quarters within a year for which they give guidance.

We find that falling short of one's own forecast has a particularly significant effect on future guidance strategy. For example, the unconditional probability for our sample that a firm guiding this year will stop guiding for next year is 30 percent. For firms consistently falling short of their own guidance this year, the probability of omitting quarterly forecasts next year increases by 8 percent relative to firms that consistently meet or beat their own guidance. Overall, we document that the economic effect of previous forecasting outcomes on forecasting strategy is comparable to many factors previously identified to affect forecasting decisions. By focusing on the multi-period nature of forecasting, we identify significant determinants of management forecast strategies that

[1] Evaluating the impact of past outcomes on forecasting behavior is complicated by the numerous possible objectives motivating the provision of guidance. The expectation-adjustment hypothesis presumes that managers have incentives to align market expectations with their private information. However, conflicts between owners and managers could create incentives to withhold information or strategically issue biased forecasts (Nagar et al. 2003; Rogers and Stocken 2005). We know of no overarching theory that encompasses all of the various motivations behind the provision of guidance.

have been largely ignored in prior literature. Our results suggest that previous forecasting outcomes are important controls that should be included in future forecasting studies.

The next section develops predictions on how quarterly guidance strategy is affected by outcomes from previously issued guidance. Sample selection criteria and descriptive statistics are discussed in Section III. Research design and main results are presented in Section IV. Robustness checks and additional evidence are presented in Section V. Section VI concludes.

II. DEVELOPMENT OF EMPIRICAL PREDICTIONS

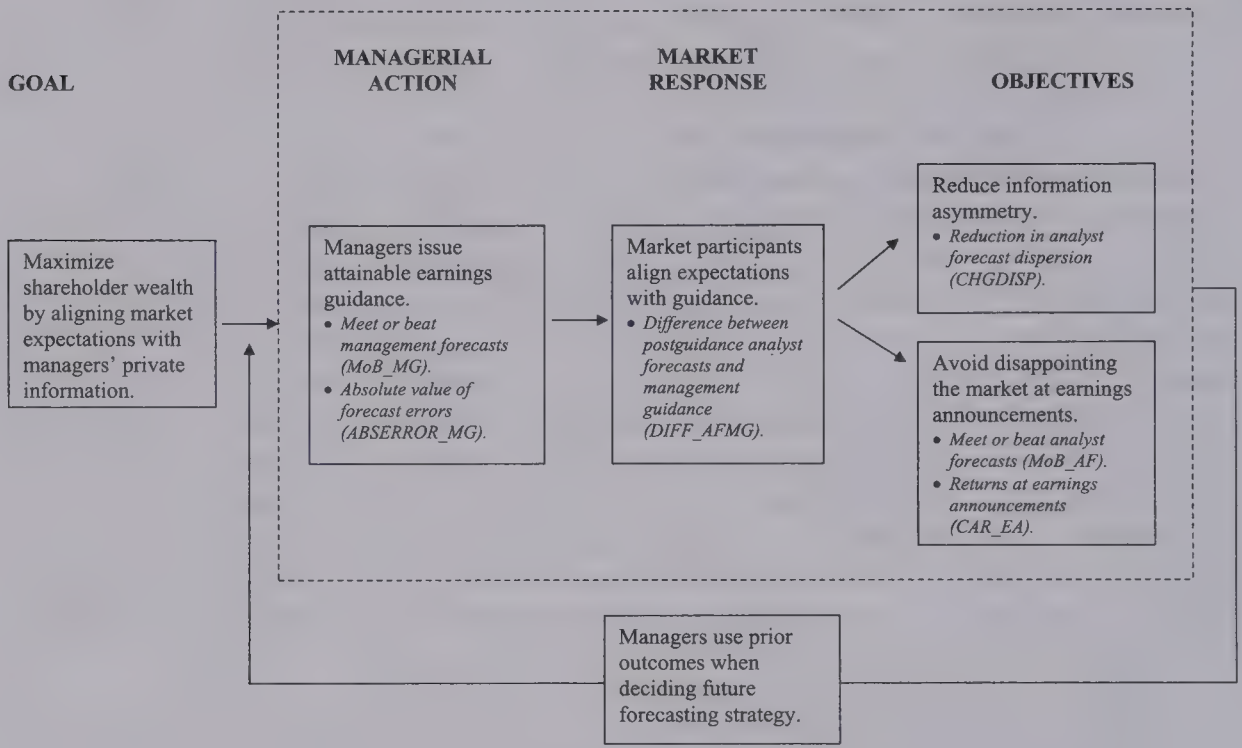
Prior research suggests a number of benefits to providing management earnings guidance, including reducing information asymmetry and cost of capital (e.g., Verrecchia 2001), reducing litigation risk (e.g., Skinner 1994, 1997), building a reputation for transparent reporting (e.g., Graham et al. 2005; Hutton and Stocken 2009), guiding analysts' forecasts toward attainable earnings targets (e.g., Cotter et al. 2006), and signaling managerial quality (Trueman 1986). Such benefits are weighed against costs to disclosure, such as exposing firms to new sources of legal liability (Kasznik 1999), proprietary costs if guidance is useful to competitors (e.g., Verrecchia 1983), and distracting management by consuming time that could otherwise be devoted to running the business (e.g., Progressive Insurance in Hutton and Weber [2001]). Prior research (e.g., King et al. 1990; Core 2001; Fields et al. 2001) suggests that firms rationally weigh expected costs and benefits when making decisions about future guidance strategy. In this section, we develop hypotheses on how outcomes from previous guidance influence managers' expectations about the costs and benefits of future guidance.

One frequently cited unifying framework for managerial objectives in providing guidance is the expectation-adjustment hypothesis (e.g., Ajinkya and Gift 1984; King et al. 1990), which states that management issues forecasts in order to align market expectations with their private information about expected future earnings. Under this view, managers issue disclosures because of the negative consequences associated with gaps between expected earnings and announced earnings, including high investor transaction costs, large stock price fluctuations, and damage to relations with analysts. In examining the relation between prior outcomes and future forecasting strategy, we adopt the expectation-adjustment framework, modified to include recent research on asymmetries in forecasting incentives. We assume that a primary goal of guidance strategy is to maximize shareholder value by aligning market expectations with managers' private information. Figure 1 summarizes our framework, dividing the effectiveness of management guidance in reaching this goal into four specific outcomes. First, managers have incentives to provide attainable guidance that can be met or beaten. Second, market participants must follow managements' guidance; otherwise, guidance is unlikely to be helpful in meeting managerial objectives. Our third and fourth outcomes relate to two specific objectives that accurate forecasts and aligning expectations with guidance can help managers to achieve: reducing information asymmetry and avoiding disappointments around earnings announcements. Each of these is discussed below.

Attainable Forecasts

Issuing attainable guidance is important to managers seeking to align market expectations with their private information. Managers have incentives to forecast accurately and would rather err on the side of being slightly pessimistic about future prospects. The benefits to forecasting accurately include signaling managerial quality (Trueman 1986), developing managerial reputation (Hutton and Stocken 2009), and increasing the credibility of future guidance (Williams 1996). Overly optimistic management guidance is particularly costly for three reasons. First, managers can be sued (under SEC Rule 10b-5) by shareholders who allege being misled into buying stock

FIGURE 1
Framework for Analyzing Forecast Outcomes
(outcome proxies in italics)



at inflated prices by the release of overly optimistic information.[2] Because most 10b-5 cases are brought after disclosures that lead to stock price declines (Francis et al. 1994; Skinner 1994), Kasznik (1999) suggests that asymmetries in legal liability rules make it more costly to fall short of one's own forecast than to exceed one's own forecast. Second, overly optimistic guidance has negative stock price consequences. For example, Soffer et al. (2000) examine the combined market reaction to preannouncement warnings and subsequent earnings announcements and find that firms that made insufficient warnings experience more negative combined market reactions than firms that preempt all bad news through preannouncement guidance. Third, recent research (e.g., Matsumoto 2002; Richardson et al. 2004; Cotter et al. 2006) has emphasized the role of earnings guidance in moving market expectations toward targets that can be met or beaten, rather than toward targets that are exactly met. This encourages managers to issue slightly pessimistic guidance.

In sum, overly optimistic guidance is costly to firms, and does not generate benefits attributed to guidance such as enhancing managerial reputation. We therefore predict that failing to meet or beat one's own forecasts will reduce the provision of guidance in the future. This could take the

[2] While the Private Securities Litigation Reform Act of 1995 introduced safe harbor provisions designed to protect the disclosure of forward-looking information, safe harbor is contingent on the inclusion of appropriate cautionary statements and requires that the issuer has no actual knowledge of falsity. Firms can still be successfully sued if the plaintiff can establish that a manager had knowledge that a forecast was false or misleading.

form of omitting all forecasts for next year or forecasting less frequently within that year.[3] In addition, firms that continue to forecast can adjust the specificity of forecasts. By issuing less specific forecasts, managers with a history of failing to meet or beat their own forecasts can decrease the probability that they will continue to fall short of their own guidance. These predictions are tested in H1:

H1: The extent to which firms meet or beat management earnings guidance is negatively related to future guidance omissions, positively related to the intra-year frequency of future guidance, and positively related to the precision of future guidance.

While we expect that costs of inaccurate disclosures arise primarily from earnings that fall short of management guidance, rather than earnings that exceed management guidance, we include absolute past forecast errors to control for management forecast accuracy in all tests of H1.

Aligning Expectations with Management's Prediction

The second guidance outcome that we consider is the extent to which post-guidance market expectations are aligned with management's predictions. This is a key element in the expectations adjustment hypothesis (e.g., Ajinkya and Gift 1984; King et al. 1990) and management's incentives to align expectations with their private information has been the focus of a number of prior studies (e.g., Kasznik and Lev 1995; Cotter et al. 2006). Managers are unlikely to obtain benefits from guiding when market participants do not follow the guidance provided. However, these managers would still bear the costs associated with providing guidance. Kothari (2001) notes that analysts' forecasts are frequently used to proxy for market expectations. Therefore, we argue that cases in which analysts do not closely follow prior guidance will discourage management from future forecasting. This is the basis for H2:

H2: The extent to which post-guidance analyst forecasts deviate from management earnings guidance is positively related to future guidance omissions and negatively related to the intra-year frequency of future guidance.

The relation between future specificity and the extent to which post-guidance analyst forecasts deviate from management guidance is less clear. When a wide gap remains between analyst expectations and guidance, managers could be unwilling to continue incurring the costs of issuing precise forecasts that do not generate the desired analyst reactions. Alternatively, such firms could be more likely to issue specific guidance if that increases the effectiveness in convincing analysts to follow guidance. We examine the relation between future specificity and the extent to which post-guidance analyst forecasts deviate from management earnings guidance but do not predict the sign of the effect.

Reducing Information Asymmetry

Theoretical work on the effects of voluntary disclosure frequently focuses on the role of information asymmetry. Public disclosures reduce informational advantages from privately acquiring information (e.g., Diamond and Verrecchia 1991). The reduction in the value and quantity of private information acquisition reduces the risk of trading against someone who has superior information. As a result, theory predicts that bid-ask spreads become narrower, while trading volume and liquidity increase (Verrecchia [2001] provides a literature review). Easley and O'Hara (2004) argue that uninformed investors avoid trading stocks with high information asymmetry,

[3] Managers could also respond to the costs of failing to meet or beat guidance by giving very pessimistic guidance. We argue that firms are unlikely to do so for two reasons. First, pessimistic forecasting could depress stock prices, leading to an increased risk of takeover or management turnover. Second, inaccurate forecasting would reduce the credibility of future disclosures.

which reduces demand for these stocks and increases the cost of capital for these firms. Easley and O'Hara (2004) suggest that firms can influence firm value and cost of capital by altering the precision and quantity of information available to investors.

Prior research suggests that one benefit to management guidance is the reduction of information asymmetry. For example, guidance has been shown to improve the accuracy of analyst forecasts (Waymire 1986), reduce the dispersion in analyst forecast (Baginski et al. 1993; Clement et al. 2003), and reduce bid-ask spreads (Coller and Yohn 1997). However, earnings guidance can increase the dispersion in beliefs when investors adopt different interpretations of news or when forecasts increase uncertainty about payoffs (e.g., Harris and Raviv 1993; Kandel and Pearson 1995).

The dispersion in analyst forecasts is a frequently used proxy for information asymmetry (e.g., Lang and Lundholm 1996; Barron and Stuerke 1998; Irani and Karamanou 2003). High dispersion denotes a low consensus among analysts about a firm's future earnings, a sign of high information asymmetry. Gebhardt et al. (2001) and Han and Manry (2000) find that firms exhibiting high forecast dispersion also exhibit high cost of capital. We use the dispersion in analyst forecasts as a proxy for information asymmetry because cost of capital is difficult to directly observe and incorporating bid-ask spread data substantially reduces our sample size.

Given that reducing information asymmetry is one main benefit obtained from issuing guidance, we argue that managers will be less willing to incur the costs of guidance when guidance has been unsuccessful at reducing information asymmetry. This is the basis for H3:

H3: The increase in the dispersion of analyst forecasts following the provision of management earnings guidance is positively related to future guidance omissions and negatively related to the intra-year frequency of future guidance.

The relation between future specificity and changes in analyst forecast dispersion is less clear. An increase in dispersion following the provision of guidance could discourage managers from issuing precise forecasts in the future, in that some firms do not want to bear the cost of precise disclosures if these are ineffective in reducing information asymmetry. Alternatively, such firms could be more likely to issue specific guidance if that increases the effectiveness in reducing analyst forecast dispersion. We examine the relation between changes in analyst forecast dispersion and future guidance specificity but do not predict the sign of the effect.

Avoiding Earnings Disappointments

The final guidance outcome that we consider is the extent to which managers avoid disappointing market participants at earnings announcements following the provision of guidance. Prior research suggests that analyst expectations are an important threshold to managers (e.g., Brown and Caylor 2005) and that earnings disappointments are typically accompanied by large stock price declines (Skinner and Sloan 2002). Using earnings guidance to preempt bad news can both reduce the probability of class-action lawsuits that follow earnings disappointments (Field et al. 2005) and reduce the magnitude of settlement amounts (Skinner 1997). Recent research suggests that an important role of earnings guidance is to move analyst expectations toward targets that can be met or beaten in order to avoid disappointing the market at earnings announcements (e.g., Matsumoto 2002; Richardson et al. 2004; Cotter et al. 2006).

Prior research also suggests that guiding and subsequently disappointing the market is more costly than simply disappointing the market. For example, Graham et al. (2005) report that interviewed CFOs consider issuing earnings guidance yet still falling short of guided analysts' forecasts to be a particularly bad outcome, as this is interpreted as evidence that the firm is poorly

managed and cannot accurately predict its own future. This is consistent with Lennox and Park (2006), who find that the market penalty for missing consensus analyst forecasts is greater for firms that issued guidance than for firms that did not issue guidance.

Given that avoiding disappointments at earnings announcements is a key benefit to providing earnings guidance, we predict firms that issued earnings guidance in the recent past but yet still disappointed the market at earnings announcements will reduce the provision of guidance in the future.[4] Again, this could take the form of omitting all forecasts for a given year or forecasting less frequently within that year. In testing this prediction, we measure earnings disappointments using both analyst forecast errors (H4a) and price reactions around earnings announcements (H4b):

H4a: The extent to which firms meet or beat guided analyst expectations is negatively related to future guidance omissions and positively related to the intra-year frequency of future guidance.

H4b: The market reaction around recent past quarterly earnings announcements that followed the provision of guidance is negatively related to future guidance omissions and positively related to the intra-year frequency of future guidance.

The relation between future specificity and the ability to avoid past earnings disappointments following the provision of guidance is less clear. Earnings disappointments following the provision of guidance could discourage managers from issuing precise forecasts in the future, in that some firms could be unwilling to bear the cost of precise disclosures if they ultimately result in earnings disappointments. Alternatively, such firms could be more likely to issue specific guidance if that increases the effectiveness in guiding expectations toward numbers that could be met or beaten. We examine the relation between earnings disappointments and guidance specificity but do not predict the sign of the effect.

III. SAMPLE SELECTION CRITERIA

We obtain earnings guidance, actual earnings, and analyst forecast information from the First Call Company Issued Guidelines (CIG) file as of March 2006. The advantage of using the CIG file is that it provides a sample size much larger than would be practical by hand-collecting data. One disadvantage is that there may be a bias in First Call's collection or reporting of guidance.[5] An additional disadvantage is that the CIG file emphasizes recent guidance, with very few observations before 1994. We obtain abnormal return data from CRSP and obtain SIC code, market value, and other accounting information from Compustat.

Table 1 summarizes the effects of imposing sample selection criteria. The March 2006 CIG file includes 43,885 quarterly management earnings forecasts. We restrict our attention to forecasts of earnings per share (EPS) denoted in U.S. dollars. Duplicate forecasts and observations in which management earnings guidance and subsequently realized earnings are not reported on a comparable basis due to a merger or accounting change are also deleted. Anilowski et al. (2007) suggest that sample sizes in the CIG file reflect progressively wider coverage of firms between 1994 and 1998, while they find that post-1998 sample sizes are more consistent. As a result, we delete forecasts for periods before 1998.[6] Our research design requires a full year of data, so we exclude

[4] This prediction is consistent with Houston et al. (2010), who find that poor performance (proxied by past experience in meeting or beating analyst expectations) helps explain why firms regularly forecasting in 2003 did not make a forecast in 2004. Our study differs from Houston et al. (2010) in that we examine the effects of a wide-range of outcomes related to previous guidance and focus on forecasting decisions made over a longer, more general sample period.

[5] A conversation with a representative from First Call confirmed that they do not intentionally follow only a subset of firms. However, Miller (2002) finds that the CIG file emphasizes quantitative forecasts, reporting a lower percentage of qualitative forecasts than is typically obtained through hand collection of disclosures.

[6] Results using samples from 1994 to 2005 are qualitatively similar and are not reported.

TABLE 1
Sample Selection

Initial sample from First Call CIG database (as of March 2006)		43,885
Currency is not in USD, duplicate observations, guidance unrelated to EPS, not comparable due to mergers or accounting changes, or without CUSIP	(1,926)	
Forecasts issued before 1998 or after 2005	(4,758)	
Forecasts made before or after the fiscal quarter	(13,233)	
Missing data to determine accuracy	(2,892)	
Multiple guidance per quarter (keeping the most recent forecast)	(2,888)	
Firm not covered by First Call in the next year	(3,030)	
Forecasts used in Table 2		15,158
Number of firm-year corresponding to these quarterly forecast		7,685
Firm-years missing data for variables in model (1)	(3,487)	
Firm-years for Table 4		4,198
Firm-years without guidance in the next year	(1,098)	
Firm-years for Table 5		3,100
Firm-years missing for variables in model (3)	(43)	
Firm-years for Table 6		3,057

forecasts related to periods after 2005. We also delete forecasts that are issued before the start of the fiscal period or after the end of the fiscal period to which they relate.[7] In the case of most range forecasts and some open-ended and qualitative forecasts, First Call reports a conversion to a point estimate, which is then used to assess *ex post* accuracy.[8] We exclude observations missing this data or missing the actual earnings number needed to calculate forecast accuracy. For firms that issue multiple forecasts for a given quarter, we keep only the last forecast made. To ensure that guidance omissions are not due to First Call dropping coverage of a firm, we require that each firm is still covered by First Call (in the company issued guidelines, summary, detailed analysts' forecasts, or actual earnings files) in the following year. These criteria yield an initial sample of 15,158 quarterly forecasts.

Table 2 presents information on forecast horizon and forecast form for this initial sample of 15,158 quarterly forecasts. We code management earnings guidance as point estimates (e.g., EPS

[7] Guidance given before the start of the fiscal period is made with significantly more uncertainty, as it relates to a longer horizon and economic activity for the prior quarter has not yet been completed. Guidance given after the fiscal period end is made with significantly less uncertainty, as economic activity for the period has already been completed. We expect that our proxies for adverse outcomes would have different effects on management behavior when they relate to very long (very short) horizon forecasts that were *ex ante* expected to be less (more) accurate.

[8] For range forecasts, First Call's conversion is typically the mid-point of the range. For maximum forecasts, First Call's conversion is typically one cent less than the maximum given. For minimum forecast, First Call's conversion is typically equal to or one cent more than the minimum given. The robustness of our results to alternate measures of accuracy is discussed in Section V.

TABLE 2

Forecast Form and Horizon

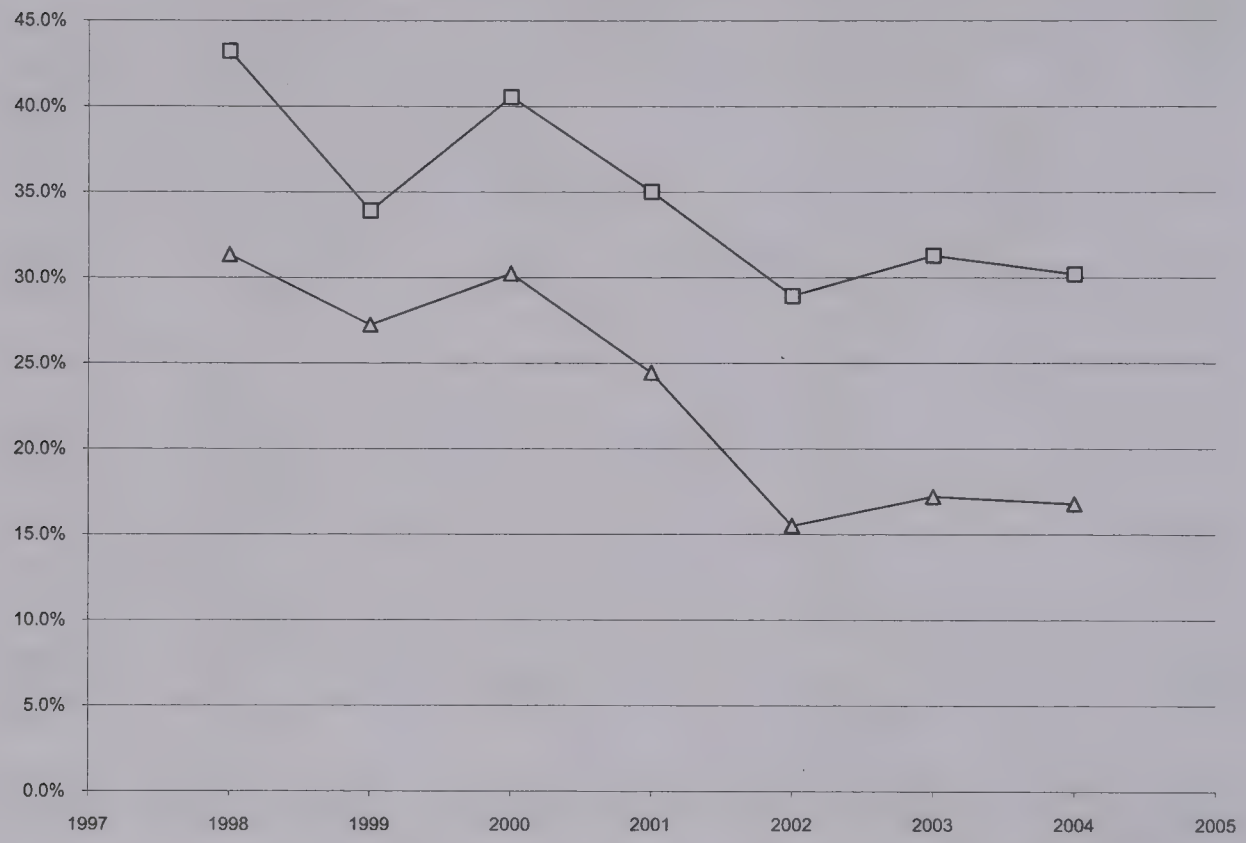
Frequency Percent Row Pct Col Pct	60–90 Days Before Fiscal Period End	30–60 Days Before Fiscal Period End	30 Days Before Fiscal Period End	Total
Point	1,185	775	1,202	3,162
	7.82%	5.11%	7.93%	20.86%
	37.48%	24.51%	38.01%	
	18.25%	22.76%	22.86%	
Range	4,976	2,409	3,461	10,846
	32.83%	15.89%	22.83%	71.55%
	45.88%	22.21%	31.91%	
	76.62%	70.75%	65.81%	
Open-ended	324	211	558	1,093
	2.14%	1.39%	3.68%	7.21%
	29.64%	19.30%	51.05%	
	4.99%	6.20%	10.61%	
Qualitative	9	10	38	57
	0.06%	0.07%	0.25%	0.38%
	15.79%	17.54%	66.67%	
	0.14%	0.29%	0.72%	
Total	6,494	3,405	5,259	15,158
	42.84%	22.46%	34.69%	100%

The sample consists of 15,158 quarterly management forecasts from 1998 to 2004 in which First Call covered the firm in the subsequent year.

is forecasted to be $1.25), range estimates (e.g., EPS is forecasted to be between $1.20 and $1.30), open-ended estimates (including maximum estimates such as "EPS will be no more than $1.30" and minimum estimates such as "EPS will be at least $1.20"), or qualitative guidance. For this sample, range estimates are the most common form of guidance (71.55 percent), followed by point estimates (20.86 percent), open-ended estimates (7.21 percent), and qualitative guidance (0.38 percent). To provide initial evidence on the effect of failing to meet or beat management guidance, we compare the frequency of optimism in forecasting for a given year conditioned on whether a firm continues to forecast. A forecast is coded as "optimistic" (*ex post* actual earnings less than management's forecast) or "neutral or pessimistic" (*ex post* actual earnings greater than or equal to management's forecast).

Figure 2 compares the sample firms that issue a quarterly forecast for the following year to those that do not. The proportion of optimistic forecasts is lower in every year of our sample period for firms forecasting for the following year than for firms that did not forecast for the following year. For example, for firms making a quarterly forecast for both 2001 and 2002, the proportion of optimistic forecasts for 2001 is 24.4 percent. In contrast, for firms making a quarterly forecast for 2001 but not for 2002, the proportion of optimistic forecasts for 2001 is 35.0 percent. Chi-squared tests (not reported) indicate that differences in the proportion of optimistic forecasts between these two groups are significant at the 0.05 level for each year from 1998 to 2004. For all sample forecasts, only 19.7 percent of the forecasts made by firms also forecasting for the next year were optimistic, while 34.6 percent of the forecasts made by other firms were

FIGURE 2
Percentage of Optimistic Forecasts for Firms Issuing Forecasts for the Next Year (Δ) and Firms Not Issuing Forecasts for the Next Year (□) from 1998 to 2004



optimistic. These results are consistent with H1, which predicts that firms issuing optimistic forecasts are more likely to stop issuing forecasts in the future than firms issuing neutral or pessimistic forecasts.

We test H1 though H4 on an annual basis, comparing outcomes in year $t-1$ to forecasting decisions for year t. It is unclear whether firms make forecasting decisions primarily on an individual quarter-by-quarter basis, or whether they adopt longer-term guidance policies. Our hypotheses should apply equally to alternate windows and specification checks presented in Section V demonstrate that results are qualitatively similar for both quarterly guidance behavior and guidance aggregated over two-year windows. Aggregating quarterly guidance into year-by-year decisions yields 7,685 firm-years. Our 7,685 firm-years include multiple observations for some firms, so we follow Petersen (2009) and report all regressions results using robust z-statistics that control for firm clustering effects. All reported p-values are two-tailed.

IV. RESEARCH DESIGN AND MAIN RESULTS

Determinants of Guidance Omissions

In this section we empirically examine the relation between prior guidance outcomes and future guidance strategy. We restrict our attention to firms that have guided in the prior year, because the focus of our research question (the decision to continue forecasting based on outcomes from prior year's guidance) is relevant only for these firms. Given our population of interest is the subset of firms that previously guided, we do not use econometric techniques to control for the

earlier decision regarding whether to guide in the prior year (e.g., the Heckman two-step procedure).[9] To test the relation between adverse outcomes and future guidance omissions, we estimate the following logistic regression:

$$Pr(OMIT_{it}) = \frac{1}{1 + e^{-\beta' X_{it-1}}} \tag{1}$$

where:

$$\begin{aligned}\beta' X_{it-1} = {} & \alpha_0 + \beta_1 MoB\_MG_{it-1} + \beta_2 ABSERROR\_MG_{it-1} + \beta_3 DIFF\_AFMG_{it-1} \\ & + \beta_4 CHGDISP_{it-1} + \beta_5 MoB\_AF_{it-1} + \beta_6 CAR\_EA_{it-1} + \beta_7 STDRET_{it-1} + \beta_8 DISP_{it-1} \\ & + \beta_9 RET_{it-1} + \beta_{10} LOSS_{it-1} + \beta_{11} INCREASE_{it-1} + \beta_{12} NUMAN_{it-1} + \beta_{13} LIT_{it-1} \\ & + \beta_{14} SIZE_{it-1} + \beta_{15} BM_{it-1} + \beta_{16} BETA_{it-1}.\end{aligned}$$

The dependent variable $OMIT_{it}$ is an indicator variable equal to 1 if firm i does not issue guidance for any quarter in year t, and 0 otherwise.[10] In examining the relation between past outcomes and future forecasting, we aggregate outcomes across quarters of the prior fiscal year into one annual measure. To test H1, we construct an indicator variable $MoB\_MG_{it-1,q}$ which takes a value of 1 if subsequently realized earnings for quarter q in year $t-1$ is greater than or equal to the earnings guidance for that quarter (i.e., if realized earnings meet or beat management's forecast), and 0 otherwise. Hypothesis 1 is tested using MoB_MG_{it-1}, which is an average of $MoB\_MG_{it-1,q}$ over each quarter for which management issued earnings guidance in the previous year. To capture the effect of absolute forecast accuracy, rather than a propensity to meet or beat management forecasts, we calculate $ABSERROR\_MG_{it-1,q}$ as the absolute value of the difference between management guidance and subsequently realized earnings for firm i in quarter q of year $t-1$, scaled by stock price at the beginning of the quarter. Our empirical tests use $ABSERROR\_MG_{it-1}$, an average of $ABSERROR\_MG_{it-1,q}$ over each quarter for which management issued guidance in the previous year.[11]

Hypothesis 2 predicts that firms whose guidance was not closely followed by analysts are less likely to continue forecasting than firms whose guidance was closely followed by analysts. To measure the extent to which analysts fail to follow management guidance, we define $DIFF\_AFMG_{it-1,q}$ as the absolute value of the difference between management guidance for quarter q in year $t-1$ and the analyst consensus forecast for that quarter as of ten days after the release of the management guidance, scaled by share price of firm i at the end of quarter $q-1$.

---

[9] The use of single-stage regressions on the subsample of interest is consistent with Wooldridge's (2001, 551) discussion of self-selection techniques. To test the sensitivity of our results to this approach, we re-estimate all analyses in this section using a Heckman two-step procedure. In re-estimating model (1), we included an initial stage regarding whether a firm issued any guidance in the prior year. In re-estimating models (2) and (3), we included an initial stage that examines whether a firm provides guidance for both the current year and the prior year. All inferences are unaffected by using this alternate estimation method.

[10] In cases where a firm makes a forecast for both quarter 4 of year $t-1$ and quarter 1 of year t, we exclude the quarter 1 forecast from our proxy for forecasting behavior in year t if the quarter 1 forecast was made prior to the earnings announcement related to the prior quarter's forecast. This insures that all forecasts included in our proxy for guidance omissions were made after the outcomes from the prior quarters are known. This affects only 67 forecasts in our sample and reported results are qualitatively unaffected by this exclusion.

[11] We also measure management forecast accuracy using the difference between absolute management forecast error and absolute forecast error based on analyst forecasts prior to management forecasts. Results using this relative measure of management forecast accuracy are qualitatively similar to those based on *ABSERROR_MG* and are not reported separately.

Hypothesis 2 is tested using $DIFF\_AFMG_{it-1}$, an average of $DIFF\_AFMG_{it-1,q}$ over each quarter for which management issued guidance in the previous year.[12]

Hypothesis 3 predicts that firms with increases in analyst dispersion following the release of past guidance are less likely to guide in the next year. To measure changes in dispersion, we define $CHGDISP_{it-1,q}$ as the standard deviation of analysts' forecasts as of ten days after the management guidance for quarter q in year $t-1$ minus the standard deviation of analysts' forecasts for that quarter as of one day before the release of the management guidance, scaled by share price of firm i at the end of quarter $q-1$. Hypothesis 3 is tested using $CHGDISP_{it-1}$, an average of $CHGDISP_{it-1,q}$ over each quarter for which management issued guidance in the previous year.

Hypothesis 4a predicts that firms that issued guidance and subsequently met or beat analysts' expectations are less likely to omit future guidance than firms that guided earnings and fell short of expectations. This prediction is tested using an indicator variable $MoB\_AF_{it-1,q}$ which takes a value of 1 if subsequently realized earnings for quarter q in year $t-1$ is greater than or equal to the analyst consensus forecast, and 0 otherwise. Analyst consensus forecasts are measured as of 12 days before the end of fiscal quarter.[13] Hypothesis 4a is tested using MoB_AF_{it-1}, an average of $MoB\_AF_{it-1,q}$ over each quarter for which management issued guidance in the previous year.

Hypothesis 4b predicts that firms that have experienced positive market reactions to earnings announcements that followed the issuance of earnings guidance are less likely to stop forecasting in the later period. To test this hypothesis, we measure unexpected returns over short windows around quarterly earnings announcements. We calculate abnormal returns $CAR\_EA_{it-1,q}$ from 12 days before the end of the fiscal quarter q in year $t-1$ to one day after the earnings announcement. Abnormal returns are calculated relative to CRSP value-weighted returns. Tests of H4b are based on CAR_EA_{it-1}, an average of $CAR\_EA_{it-1,q}$ over each quarter for which management issued earnings guidance in the previous year.

Model (1) includes two controls for prior period uncertainty: $STDRET_{it-1}$ (calculated as the standard deviation of raw daily stock returns for firm i over year t) and $DISP_{it-1}$ (calculated as the average of analyst forecast dispersion at the start of the quarter for all quarters of firm i in year $t-1$). Miller (2002) finds that firms with sustained earnings increases contemporaneously increase the level of voluntary disclosure, while Chen et al. (2007) find that poorly performing firms are more likely to renounce the provision of earnings guidance. To control for prior year performance, we include three variables taken from Chen et al. (2007). RET_{it-1} is calculated as abnormal returns (relative to CRSP value-weighted returns) for firm i cumulated over year $t-1$. $LOSS_{it-1}$ is calculated as the percentage of quarters for firm i with losses in year $t-1$. $INCREASE_{it-1}$ is calculated as the percentage of quarters for firm i with earnings increases relative to a seasonal random walk model for year $t-1$. *LOSS* and *INCREASE* are based on Compustat quarterly income before extraordinary items. Based on the results of Chen et al. (2007), $LOSS_{it-1}$ is expected to be positively related to the decision to stop forecasting, while RET_{it-1} and $INCREASE_{it-1}$ are expected to be negatively related.

---

[12] As a specification check we re-estimate model (1) replacing *DIFF_AFMG* with *RESPONSE*, calculated as the change in analysts' forecasts around management guidance, scaled by the initial difference between management guidance and pre-guidance analysts' forecasts. To minimize the impact of small denominators, we exclude forecasts where the absolute value of the initial difference between management guidance and pre-guidance analysts' forecasts is less than or equal to \$0.03. When replacing $DIFF\_AFMB_{it-1}$ with $RESPONSE_{it-1}$ in model (1), the coefficient on *RESPONSE* is significantly negative in the restricted model ($\beta = -0.3350$, p = 0.02) and marginally significant in the full model ($\beta = -0.2314$, p = 0.10). Results for *RESPONSE* are insignificant when observations with all denominators (other than zero) are included.

[13] We measure earnings surprises as of 12 days before the fiscal quarter end to be consistent with Skinner and Sloan (2002), who find that much of the market reaction to earnings news occurs prior to a three-day window centered on earnings announcements. Our return window for price reactions to earnings announcements also corresponds to Skinner and Sloan (2002).

Model (1) also includes a number of additional control variables that prior research suggests are associated with forecasting decisions. To control for litigation environment, we follow Francis et al. (1994) by using a qualitative variable LIT_{it-1} to indicate membership for firm i in year $t-1$ in biotechnology (SIC codes 2833–2836, SIC codes 8731–8734), computer (SIC codes 3570–3577, SIC codes 7370–7374), electronics (SIC codes 3600–3674), and retail (SIC codes 5200–5961) industries. We define $NUMAN_{it-1}$ as the average number of analysts following firm i as of the beginning of the quarter for all quarters in year $t-1$. A number of prior papers (e.g., Lang and Lundholm 1993) document a positive association between analyst following and disclosure quality. $SIZE_{it-1}$ is measured as the natural logarithm of total assets for firm i at the end of year $t-1$. Based on prior research (e.g., Kasznik and Lev 1995), firm size is expected to be positively related to disclosure frequency. BM_{it-1} is the ratio of book value to market value of equity for firm i at the end year $t-1$. Market to book (the inverse of our measure) is used by Bamber and Cheon (1998) to proxy for proprietary costs. $BETA_{it-1}$ is calculated as the coefficient from estimating Sharpe's (1964) market model over year $t-1$. Ajinkya et al. (2005) report a positive relation between systematic risk and disclosure frequency. The data requirements on control variables yield a sample size of 4,198 firm-year observations, of which 3,100 issue at least one forecast in the subsequent year and 1,098 do not.

Table 3 reports Spearman and Pearson correlation tables for $OMIT_{it}$ and our treatment variables. These correlations are consistent with our H1 through H4b. In particular, future guidance omissions ($OMIT_{it}$) are negatively correlated with meeting or beating one's forecast (MoB_MG_{it-1}), positively correlated with the extent to which post-guidance analyst forecasts differ from management guidance ($DIFF\_AFMG_{it-1}$), positively correlated with increases in analyst forecast dispersion following guidance ($CHGDISP_{it-1}$), and negatively correlated with our proxies for avoiding earnings disappointments (MoB_AF_{it-1} and CAR_EA_{it-1}). Each of these relations is statistically significant using both Pearson and Spearman correlations (one exception is the Spearman correlation between $OMIT_{it}$ and $CHGDISP_{it-1}$). Table 3 also suggests that our treatment variables are significantly correlated with each other. Therefore, regression results are presented for both a full model and by including the variables used to test H1 through H4b individually. We also report specification checks using principal component analysis.

Table 4 reports results for model (1). Marginal effects for each independent variable in the full model are calculated as the change in the probability of omitting guidance for a given year that results from moving from the first to third quartile of each independent variable's distribution. Hypothesis 1 predicts a negative relation between the decision to stop forecasting and the ability to meet or beat management's past earnings guidance. Consistent with H1, we find a significant negative relation between the probability of omitting forecasts and *MoB_MG* in both the reduced model ($\beta_1 = -0.4860$, $p < 0.01$) and the full model ($\beta_1 = -0.4252$, $p < 0.01$). In contrast, *ABSERROR_MG*, which proxies for the absolute forecast error of past management guidance is insignificant in both the reduced ($\beta_2 = 9.2583$, $p = 0.11$) and full model ($\beta_2 = 0.6403$, $p > 0.50$). Results therefore suggest that future forecasting decisions are affected by the ability to meet or beat management guidance, rather than the unsigned accuracy of previous guidance.

Hypothesis 2 predicts a positive relation between forecast omissions and the extent to which post-guidance analyst forecasts deviate from prior guidance. Consistent with H2, we find a significant positive relation between the probability of omitting forecasts and *DIFF_AFMG* in both the reduced ($\beta_3 = 43.8820$, $p < 0.01$) and full model ($\beta_3 = 33.1874$, $p = 0.01$).

Hypothesis 3 predicts that the probability of omitting forecasts is increasing in the change in dispersion of analyst forecasts following the release of previous earnings guidance. Consistent with this predication, we find a positive relation between the probability of omitting forecasts and *CHGDISP* that is significant in the reduced model ($\beta_4 = 54.4515$, $p = 0.01$) and marginally significant in the full model ($\beta_4 = 42.2345$, $p = 0.06$).

TABLE 3

Spearman/Pearson Correlation Matrix

	$OMIT_{it}$	MoB_MG_{it-1}	$ABSERROR\_MG_{it-1}$	$DIFF\_AFMG_{it-1}$	CHG_DISP_{it-1}	MoB_AF_{it-1}	CAR_EA_{it-1}
$OMIT_{it}$		−0.1563	0.1318	0.1277	0.0211	−0.1269	−0.1310
		<0.0001	<0.0001	<0.0001	0.1728	<0.0001	<0.0001
MoB_MG_{it-1}	−0.1014		−0.2751	−0.1872	0.0043	0.6102	0.2616
	<0.0001		<0.0001	<0.0001	0.7789	<0.0001	<0.0001
$ABSERROR\_MG_{it-1}$	0.0419	−0.2533		0.5208	−0.0413	−0.2162	−0.1140
	0.0066	<0.0001		<0.0001	0.0075	<0.0001	<0.0001
$DIFF\_AFMG_{it-1}$	0.0599	−0.2116	0.3726		−0.0098	−0.1545	−0.0246
	0.0001	<0.0001	<0.0001		0.5265	<0.0001	0.1108
$CHGDISP_{it-1}$	0.0588	0.0175	−0.0691	−0.0527		0.0016	−0.0332
	0.0001	0.2575	<0.0001	0.0006		0.9186	0.0313
MoB_AF_{it-1}	−0.0683	0.6187	−0.1919	−0.1687	0.0213		0.3049
	<0.0001	<0.0001	<0.0001	<0.0001	0.1679		<0.0001
CAR_EA_{it-1}	−0.1286	0.2622	−0.0042	−0.0298	−0.0264	0.2924	
	<0.0001	<0.0001	0.7843	0.0532	0.0873	<0.0001	

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under two-tailed tests.

The sample consists of 4,198 firm years between 1998 and 2005 with available data for all variables.

Spearman correlations are reported above the diagonal, and Pearson correlations are reported below.

Variable Definitions:

$OMIT_{it}$ = indicator variable equal to 1 if firm i does not issue guidance for any quarter in year t, and 0 otherwise;

MoB_MG_{it-1} = average over the previous year of an indicator variable equal to 1 if realized quarterly EPS is greater than or equal to management's forecast, and 0 otherwise;

$ABSERROR\_MG_{it-1}$ = average over the previous year of the absolute management forecast errors (scaled by stock price at the beginning of the quarter);

$DIFF\_AFMG_{it-1}$ = average over the previous year of the absolute difference between management forecasts and analyst consensus forecasts after management forecasts (scaled by stock price at the beginning of the quarter);

$CHGDISP_{it-1}$ = average over the previous year of the standard deviation of analysts' forecasts as of 10 days after management guidance minus the standard deviation of analysts' forecasts as of 1 day before the release of the management guidance (scaled by stock price at the beginning of the quarter);

MoB_AF_{it-1} = average over the previous year of an indicator variable equal to 1 if realized quarterly EPS is greater than or equal to the analyst consensus forecast for a quarter for which management issued guidance, and 0 otherwise; and

TABLE 3 (continued)

CAR_EA_{it-1} = average over the previous year of abnormal returns from 12 days before the end of the quarter to one day after each earnings announcement for all quarters for which management issued guidance.

TABLE 4

Logistic Regression on the Probability of Omitting Guidance for Year *t*

Variable	Predicted Sign	Coefficient Estimates (robust z-statistics)					Marginal Effects
MoB_MG_{it-1}	−	−0.4860***				−0.4252***	−0.0411
		(−4.44)				(−3.32)	
$ABSERROR\_MG_{it-1}$	+	9.2583				0.6403	0.0003
		(1.58)				(0.09)	
$DIFF\_AFMG_{it-1}$	+		43.8820***			33.1874**	0.0052
			(3.85)			(2.46)	
$CHGDISP_{it-1}$	+			54.4515**		42.2345*	0.0050
				(2.52)		(1.91)	
MoB_AF_{it-1}	−				−0.2782**	0.0202	0.0013
					(−2.33)	(0.14)	
CAR_EA_{it-1}	−				−0.9103***	−0.8559***	−0.0231
					(−3.03)	(−2.85)	
$STDRET_{it-1}$		20.6699***	21.6988***	22.0412***	22.0073***	20.8230***	0.0803
		(6.51)	(6.88)	(7.00)	(6.91)	(6.50)	
$DISP_{it-1}$		−2.1728	−9.1017	21.1023	11.0200	6.5076	0.0021
		(−0.13)	(−0.58)	(1.40)	(0.76)	(0.39)	
RET_{it-1}		−0.3603***	−0.4225***	−0.4066***	−0.2977***	−0.2885***	−0.0318
		(−3.27)	(−3.67)	(−3.60)	(−2.69)	(−2.66)	
$LOSS_{it-1}$		0.3792**	0.3959**	0.4134***	0.3261**	0.3299**	0.0159
		(2.42)	(2.52)	(2.64)	(2.07)	(2.09)	
$INCREASE_{it-1}$		−0.0593	−0.1524	−0.1441	−0.0799	−0.0480	−0.0046
		(−0.51)	(−1.35)	(−1.28)	(−0.69)	(−0.41)	
$NUMAN_{it-1}$		−0.0486***	−0.0501***	−0.0509***	−0.0505***	−0.0492***	−0.0665
		(−4.68)	(−4.88)	(−4.91)	(−4.86)	(−4.75)	
LIT_{it-1}		−0.3894***	−0.3782***	−0.3903***	−0.3877***	−0.3849***	−0.0743
		(−4.01)	(−3.88)	−4.03)	(−4.00)	(−3.96)	

(continued on next page)

TABLE 4 (continued)

Variable	Predicted Sign	Coefficient Estimates (robust z-statistics)					Marginal Effects
$SIZE_{it-1}$		0.1120***	0.1146***	0.1147***	0.1192***	0.1148***	0.0516
		(3.49)	(3.57)	(3.56)	(3.69)	(3.55)	
BM_{it-1}		0.0249	0.0345	0.0406	0.0129	−0.0058	−0.0005
		(0.26)	(0.36)	(0.43)	(0.14)	(−0.06)	
$BETA_{it-1}$		−0.4406***	−0.4522***	−0.4675***	−0.4738***	−0.4287***	−0.0619
		(−5.02)	(−5.19)	(−5.34)	(−5.38)	(−4.86)	
Intercept		−1.3052***	−1.6625***	−1.6436***	−1.4703***	−1.4042***	
		(−4.69)	(−6.07)	(−5.98)	(−5.23)	(−4.96)	
Model Chi-Square		253.54	237.27	237.44	247.84	272.90	

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under two-tailed tests.

Robust z-statistics are based on standard error estimates that control for firm clustering effects.

The sample consists of 4,198 firm years between 1998 and 2005 with available data for all variables. The dependent variable, $OMIT_{it}$, is an indicator variable equal to 1 if firm i does not issue guidance for any quarter in year t, and 0 otherwise.

Variable Definitions:

$STDRET_{it-1}$ = standard deviation of daily stock returns over the previous year;
$DISP_{it-1}$ = average over the previous year of analyst forecast dispersion at the beginning of each quarter (scaled by stock price at the beginning of the quarter);
RET_{it-1} = abnormal returns cumulated over the previous year;
$LOSS_{it-1}$ = percentage of quarters with losses in the previous year;
$INCREASE_{it-1}$ = percentage of quarters with earnings increases relative to a seasonal random walk model for the previous year;
$NUMAN_{it-1}$ = average over the previous year of the number of analysts following the firm at the beginning of each quarter;
LIT_{it-1} = indicator variable equal to 1 if the firm is in the biotechnology, computer, electronics, or retailing sector, and 0 otherwise;
$SIZE_{it-1}$ = natural logarithm of total assets;
BM_{it-1} = ratio of book value to market value of equity; and
$BETA_{it-1}$ = slope coefficient from estimating Sharpe's (1964) market model using daily return data from the previous year.

All other variables are defined in Table 3.

Hypothesis 4a predicts a negative relation between the decision to stop forecasting and past ability to meet or beat guided analyst earnings forecasts. Consistent with H4a, we find a significant negative relation between the probability of omitting forecasts and *MoB_AF* in the reduced model ($\beta_5 = -0.2782$, $p = 0.02$). However, the relation between the probability of omitting forecasts and *MoB_AF* is insignificant ($\beta_5 = 0.0202$, $p > 0.50$) in the full model. H4b predicts a negative relation between the probability of omitting forecasts and prior price responses to earnings news. As predicted, we find a negative and statistically significant relation between the probability of omitting forecasts and *CAR_EA* in both the reduced and full models ($\beta_6 = -0.9103$, $p < 0.01$; $\beta_6 = -0.8559$, $p < 0.01$, respectively).[14]

Among our control variables, *STDRET, RET*, *LOSS*, *NUMAN*, *LIT*, *SIZE*, and *BETA* are statistically significant at the 0.05 level. Large firms and firms that experienced a high standard deviation of returns or a high proportion of quarterly losses over the past year are more likely to omit forecasts. Firms that experienced positive returns over the past year, firms with high analyst following, firms operating in industries particularly at risk for security litigation, and firms with high systematic risk are less likely to omit forecasts.

Overall, results from model (1) are consistent with our hypotheses that omissions in the provision of guidance are influenced by the outcomes from previously issued guidance. Among the treatment variables, only *MoB_AF* is not consistently related to future guidance omissions. Marginal effects for our treatment variables suggest that variation in *MoB_MG* has the largest overall effect. In the full model, moving from the first to the third quartile of *MoB_MG* reduces the probability of omitting forecasts by approximately 4 percent. The marginal effect when moving from *MoB_MG* of 0 to *MoB_MG* of 1 (not reported in the tables) is 8 percent.[15,16] These effects appear economically significant, particularly when compared to 30 percent, which is the baseline probability that a firm forecasting in year $t-1$ omits forecasts in year t.

Determinants of Intra-Year Forecasting Frequency

The previous section examines the extent to which omissions in the provision of earnings guidance are influenced by the outcomes from previously issued earnings guidance. We next examine the extent to which firms issuing forecasts for a given year condition forecast frequency on outcomes from previously issued guidance. Restricting our sample to observations in which a

---

[14] As Table 3 shows, many of the treatment variables are correlated, which raises concerns about the relevance of individual outcomes. We examine the robustness of results for model (1) by performing principal component analysis. By examining the scree plot and eigenvalues, we identify two factors. The first factor is primarily correlated with *MoB_MG*, *MoB_AF*, and *CAR_EA*, capturing meeting or beating key earnings benchmarks. The second factor is primarily correlated with *ABSERROR_MG* and *DIFF_AFMG*, capturing absolute management forecast errors and a lack of analyst response to guidance. We re-estimate model (1), replacing *MoB_MG*, *ABSERROR_MG*, *MoB_AF*, *CAR_EA*, and *DIFF_AFMG* with these two factors. We find a significantly negative coefficient on the factor related to falling short of earnings benchmarks ($\beta_1 = -0.2165$, $p < 0.01$) and a significantly positive coefficient on the factor related to absolute management forecast errors and the extent to which analysts deviate management guidance ($\beta_2 = 0.1715$, $p < 0.01$). The coefficient on *CHGDISP* remains significantly positive ($\beta_3 = 43.4858$, $p = 0.03$). Overall, results utilizing principal component analysis are consistent with the results of model (1).

[15] Prior research (e.g., Kasznik and Lev 1995) argues that managers give guidance to minimize the costs of large gaps between expected and announced earnings. As a specification check, we re-estimate model (1) using two proxies for such a gap. We calculate $EXPGAP_{itq}$ as the absolute value of the difference for firm i in year t between realized quarterly EPS for quarter q and the consensus analyst forecasts for quarter q as of the beginning of the quarter, scaled by beginning of quarter stock price. Our empirical tests use an average of $EXPGAP_{itq}$ calculated over each quarter for year t. We calculate $EXPDUM_{it}$ as an indicator variable equal to 1 for firm i in year t if the consensus analyst forecast error as of the beginning of any quarter in year t is greater than 1 percent of firm value. Results including either proxy are qualitatively similar to those of Table 4.

[16] We re-estimate model (1) after excluding observations where data from Execucomp indicates that CEO turnover occurs between year $t-1$ and year t. These results are qualitatively similar to results in Table 4.

firm makes at least one quarterly forecast for two successive years reduces our sample size to 3,100 observations.[17] We define frequency as the number of quarters within that year for which the firm makes a forecast and estimate the following ordered probit model:

$$\begin{aligned} Y_{it}^{*} = {} & \alpha_0 + \beta_1 MoB\_MG_{it-1} + \beta_2 ABSERROR\_MG_{it-1} + \beta_3 DIFF\_AFMG_{it-1} + \beta_4 CHGDISP_{it-1} \\ & + \beta_5 MoB\_AF_{it-1} + \beta_6 CAR\_EA_{it-1} + \beta_7 STDRET_{it-1} + \beta_8 DISP_{it-1} + \beta_9 RET_{it-1} \\ & + \beta_{10} LOSS_{it-1} + \beta_{11} INCREASE_{it-1} + \beta_{12} NUMAN_{it-1} + \beta_{13} LIT_{it-1} + \beta_{14} SIZE_{it-1} \\ & + \beta_{15} BM_{it-1} + \beta_{16} BETA_{it-1} + \varepsilon_{it}, \end{aligned} \tag{2}$$

under the assumption that for an underlying construct k:

$$NUMQTRS_{it} = 1 \text{ if } Y_{it}^{*} < k_1,$$

$$NUMQTRS_{it} = 2 \text{ if } k_1 \leq Y_{it}^{*} < k_2,$$

$$NUMQTRS_{it} = 3 \text{ if } k_2 \leq Y_{it}^{*} < k_3,$$

$$NUMQTRS_{it} = 4 \text{ if } k_3 \leq Y_{it}^{*},$$

where $NUMQTRS_{it}$ takes a value of 1 to 4 if firm i forecasts for one to four quarters in year t, respectively. All other variables are as defined in model (1).

Table 5 presents results from estimating the ordered probit model (2). As with model (1), we report results for both the full model and restricted models that separately include our treatment variables. Consistent with H1, we find a positive association between *NUMQTRS* and *MoB_MG* that is significant in the restricted model ($\beta_1 = 0.2464$, $p < 0.01$) but insignificant in the full model ($\beta_1 = 0.1006$, $p = 0.23$). *ABSERROR_MG*, which proxies for the absolute forecast error of past management guidance, is insignificant in both the reduced model ($\beta_2 = -4.0562$, $p = 0.32$) and the full model ($\beta_2 = 3.9973$, $p = 0.41$). Consistent with H2, we find a significantly negative relation between *NUMQTRS* and *DIFF_AFMG* in both the reduced ($\beta_3 = -29.0805$, $p < 0.01$) and full model ($\beta_3 = -28.4273$, $p < 0.01$). Consistent with H3, we find a significantly negative relation between *NUMQTRS* and *CHGDISP* in both the reduced ($\beta_4 = -47.3454$, $p < 0.01$) and the full model ($\beta_4 = -38.8935$, $p < 0.01$). Consistent with H4a, we find a significantly positive relation between forecast frequency and *MoB_AF* in the reduced model ($\beta_5 = 0.3109$, $p < 0.01$) and the full model ($\beta_5 = 0.2194$, $p = 0.01$). Hypothesis 4b is not supported, as we find a positive, but insignificant, relation between *NUMQTRS* and *CAR_EA* in the reduced model ($\beta_6 = 0.1405$, $p = 0.36$) and the full model ($\beta_6 = 0.1348$, $p = 0.39$).[18]

Overall, results from model (2) are consistent with our hypotheses that forecasting frequency is influenced by certain outcomes from previously issued earnings guidance. Conditional on issuing a quarterly forecast for a given year, we find that the number of quarters for which a firm forecasts is decreasing in the extent to which analyst forecasts differed from management forecasts

[17] This specification suffers from a self-selection bias because a firm must decide to forecast before forecast frequency can be observed. A standard approach to self-selection bias would be to utilize a Heckman two-step procedure, which in this case is complicated by the fact that the second stage is an ordered probit model. Therefore, we report results from a one-stage ordered probit model and conduct two specification checks. First, all results using the Heckman two-step procedure are qualitatively similar to those for model (2). Second, allowing *NUMQTRS* to take a value of 0 if the firm did not forecast at all also yields qualitatively similar results.

[18] Re-estimating model (2) using principal component analysis yields similar results. We find a significantly positive coefficient on the factor related to meeting or beating earnings benchmarks ($\beta_1 = 0.1325$, $p < 0.01$) and a significantly negative coefficient on the factor related to absolute management forecast errors and a lack of analyst response ($\beta_2 = -0.0835$, $p = 0.03$). The coefficient on *CHGDISP* remains significantly negative ($\beta_3 = -50.1520$, $p < 0.01$).

TABLE 5

Ordered Probit Regressions on the Number of Quarters for Which a Company Issues a Forecast for Year *t*

Variable	Predicted Sign	Coefficient Estimates (robust z-statistics)				
MoB_MG_{it-1}	+	0.2464***				0.1006
		(3.69)				(1.21)
$ABSERROR\_MG_{it-1}$	−	−4.0562				3.9973
		(−0.99)				(0.82)
$DIFF\_AFMG_{it-1}$	−		−29.0805***			−28.4273***
			(−3.86)			(−3.14)
$CHGDISP_{it-1}$	−			−47.3454***		−38.8935***
				(−3.67)		(−2.99)
MoB_AF_{it-1}	+				0.3109***	0.2194**
					(4.27)	(2.44)
CAR_EA_{it-1}	+				0.1405	0.1348
					(0.91)	(0.86)
$STDRET_{it-1}$		−22.2541***	−22.5880***	−22.9496***	−23.1068***	−22.8245***
		(−10.62)	(−10.81)	(−11.01)	(−11.00)	(−10.78)
$DISP_{it-1}$		15.1986	20.7184**	−2.7714	11.1277	6.7638
		(1.40)	(2.06)	(−0.29)	(1.17)	(0.60)
RET_{it-1}		0.0677**	0.0814**	0.0767**	0.0596*	0.0609*
		(2.06)	(2.25)	(2.23)	(1.83)	(1.85)
$LOSS_{it-1}$		−0.0500	−0.0673	−0.0652	−0.0164	−0.0283
		(−0.55)	(−0.74)	(−0.72)	(−0.18)	(−0.31)
$INCREASE_{it-1}$		0.1280**	0.1722***	0.1616***	0.1121*	0.1034
		(2.02)	(2.81)	(2.65)	(1.79)	(1.62)
$NUMAN_{it-1}$		0.0142**	0.0146***	0.0153***	0.0144**	0.0139**
		(2.52)	(2.62)	(2.74)	(2.55)	(2.47)
LIT_{it-1}		0.1291**	0.1238**	0.1271**	0.1256**	0.1257**
		(2.17)	(2.08)	(2.14)	(2.11)	(2.12)

(continued on next page)

TABLE 5 (continued)

Variable	Predicted Sign	Coefficient Estimates (robust z-statistics)				
$SIZE_{it-1}$		−0.0769***	−0.0778***	−0.0777***	−0.0803***	−0.0785***
		(−3.86)	(−3.91)	(−3.89)	(−4.03)	(−3.93)
BM_{it-1}		0.0219	0.0163	0.0145	0.0318	0.0464
		(0.34)	(0.25)	(0.22)	(0.49)	(0.71)
$BETA_{it-1}$		0.4159***	0.4200***	0.4207***	0.4284***	0.4128***
		(8.61)	(8.70)	(8.71)	(8.87)	(8.49)
Intercept		−1.1474***	−1.3376***	−1.3422***	−1.1218***	−1.1268***
		(−6.24)	(−7.50)	(−7.49)	(−6.18)	(−6.12)
Intercept 2		0.5472***	0.5475***	0.5474***	0.5480***	0.5510***
		(2.99)	(3.08)	(3.07)	(3.04)	(3.01)
Intercept 3		1.1397***	1.1397***	1.1393***	1.1405***	1.1466***
		(6.27)	(6.46)	(6.43)	(6.36)	(6.31)
Model Chi-Square		234.41	235.14	236.72	237.85	272.13

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under two-tailed tests.

Robust z-statistics are based on standard error estimates that control for firm clustering effects.

The sample consists of 3,100 firm years between 1998 and 2005 for firms that issue at least one forecast for both the preceding and current year and have data for all independent variables. The dependent variable equals 4 if a firm issues forecasts for all four quarters in year *t*, 3 if a firm issues forecasts for three quarters, 2 if a firm issues forecasts for two quarters, and 1 if a firm issues a forecast for one quarter.

All variables are defined in Table 3 and Table 4.

following the provision of guidance, decreasing in the change in dispersion that followed the issuance of earnings guidance for the previous year, and increasing in firm's past ability to meet or beat guided analysts' forecasts. While reduced model results suggest that forecast frequency is increasing in firm's past ability to meet or beat their own forecasts for the prior year, this effect is insignificant in the full model.

Determinants of Forecast Specificity

We next examine the extent to which past adverse outcomes from forecasts affect the specificity of future forecasts. To proxy for specificity, we construct a variable $FORM_{it,q}$ which equals 3 if the last forecast made for quarter q in year t is a point estimate, 2 for a range estimate, 1 for an open-ended estimate (maximum or minimum), and 0 for a qualitative statement. $FORM_{it,q}$ is therefore increasing in the precision or specificity of guidance. Empirical tests use $FORM_{it}$, an average of $FORM_{it,q}$ over each quarter for which management issued earnings guidance during the year.

To test the effect of outcomes from previous guidance on forecast form we estimate the following linear regression equation:

$$\begin{aligned} FORM_{it} = {} & \alpha_0 + \beta_1 MoB\_MG_{it-1} + \beta_2 ABSERROR\_MG_{it-1} + \beta_3 DIFF\_AFMG_{it-1} \\ & + \beta_4 CHGDISP_{it-1} + \beta_5 MoB\_AF_{it-1} + \beta_6 CAR\_EA_{it-1} + \beta_7 DAYS_{it} + \beta_8 STDRET_{it-1} \\ & + \beta_9 CARSIGN_{it} + \beta_{10} NUMAN_{it-1} + \beta_{11} SIZE_{it-1} + \beta_{12} FORM_{it-1} + \varepsilon_{it}. \end{aligned} \quad (3)$$

Following Baginski and Hassell (1997), model (3) controls for the effect of forecast horizon, the standard deviation of raw daily stock returns, the sign of the price reaction to the forecast, analyst following, and firm size on management forecast precision. $STDRET_{it-1}$, $NUMAN_{it-1}$, and $SIZE_{it-1}$ are each defined earlier for model (1). $DAYS_{it}$ is calculated as the average number of days between the fiscal period end and release of guidance for each quarter in year t for which management issued guidance. $CARSIGN_{it,q}$ is an indicator variable equal to 1 if the three-day abnormal returns around a guidance date for firm i in quarter q of year t is positive, and 0 otherwise. Our empirical tests use $CARSIGN_{it}$, an average of $CARSIGN_{it,q}$ over each quarter in year t for which management issued guidance. $FORM_{it-1}$ is included to control for past behavior regarding forecast specificity. Data requirements to calculate these additional control variables reduce our sample size from 3,100 to 3,057 observations.

To control for selection bias, we estimate model (3) using a Heckman two-step procedure. In the first step, we use the independent variables of model (1) to model the initial decision of whether to forecast, using all firm-years in our sample. Estimated parameters from this probit model are then used to calculate the inverse Mills ratio, which is included as an additional independent variable in the estimation of model (3). Table 6 presents results from estimating model (3) over observations from 1998 to 2005. While we find a marginally significant positive relation between *FORM* and *MoB_MG* in the reduced model ($\beta_1 = 0.0665$, p = 0.08), this relation is insignificant in the full model ($\beta_1 = 0.0732$, p = 0.12). The coefficient on *DIFF_AFMG* is insignificant in the reduced model ($\beta_3 = -7.3427$, p = 0.14) and significantly negative in the full model ($\beta_3 = -12.7623$, p = 0.02). Full model results therefore suggest that forecast specificity is decreasing in the deviation of post-guidance analyst forecasts from prior management guidance. The coefficient on *CAR_EA* is insignificant in the reduced model ($\beta_6 = 0.1227$, p = 0.20), but marginally significant in the full model ($\beta_6 = 0.1846$, p = 0.06). Full model results therefore suggest that forecast specificity is increasing in the market reaction to earnings announcements

TABLE 6

Linear Regression on Average Forecast Specificity for Year *t*

Variable	Predicted Sign	Coefficient Estimates (robust z-statistics)				
MoB_MG_{it-1}	+	0.0665*				0.0732
		(1.76)				(1.55)
$ABSERROR\_MG_{it-1}$		0.0631				0.1862
		(0.37)				(0.99)
$DIFF\_AFMG_{it-1}$			−7.3427			−12.7623**
			(−1.49)			(−2.38)
$CHGDISP_{it-1}$				−4.9616		−7.3221
				(−0.80)		(−1.16)
MoB_AF_{it-1}					0.0383	0.0036
					(0.97)	(0.07)
CAR_EA_{it-1}					0.1227	0.1846*
					(1.27)	(1.86)
$DAYS_{it}$		0.0328*	0.0320*	0.0315*	0.0334*	0.0331*
		(1.78)	(1.75)	(1.72)	(1.82)	(1.80)
$STDRET_{it-1}$		−1.1632	−1.0594	−1.1022	−1.2394	−1.7745*
		(−1.19)	(−1.09)	(−1.12)	(−1.25)	(−1.72)
$CARSIGN_{it}$		0.0542*	0.0556*	0.0547*	0.0533	0.0513
		(1.68)	(1.72)	(1.70)	(1.64)	(1.59)
$NUMAN_{it-1}$		0.0064**	0.0057**	0.0055**	0.0069**	0.0093***
		(2.23)	(2.10)	(2.03)	(2.35)	(2.99)
$SIZE_{it-1}$		−0.0290***	−0.0279***	−0.0265***	−0.0296***	−0.0367***
		(−2.96)	(−2.87)	(−2.77)	(−2.98)	(−3.58)
$FORM_{it-1}$		0.2372***	0.2342***	0.2395***	0.2386***	0.2301***
		(8.57)	(8.44)	(8.65)	(8.61)	(8.28)
$INVMILLS_{it}$		−0.1813*	−0.1991*	−0.2416**	−0.1567	−0.0063
		(−1.65)	(−1.93)	(−2.54)	(−1.47)	(−0.05)

(continued on next page)

TABLE 6 (continued)

Variable	Predicted Sign	Coefficient Estimates (robust z-statistics)				
Intercept		1.6377***	1.7098***	1.7026***	1.6489***	1.6268***
		(13.78)	(15.01)	(14.88)	(13.92)	(13.66)
Adj. R^2		0.0879	0.0879	0.0869	0.0879	0.0919

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under two-tailed tests.

Robust z-statistics are based on standard error estimates that control for firm clustering effects.

The sample consists of 3,057 firm years between 1998 and 2005 for firms that issue at least one forecast for both the preceding and current year and have data for all independent variables. The dependent variable is the average over year *t* of a variable that takes a value of 3 if a quarter's forecast is a point estimate, 2 if a quarter's forecast is a range estimate, 1 if a quarter's forecast is an open-ended estimate, and 0 if a quarter's forecast is a qualitative statement.

Variable Definitions:

$DAYS_{it}$ = average of the day difference between fiscal period end and management forecast date;
$CARSIGN_{it}$ = average of an indicator variable equal to 1 if three-day abnormal returns around management forecast dates are positive, and 0 otherwise;
$FORM_{it-1}$ = average over year $t-1$ of a variable that takes a value of 3 if a quarter's forecast is a point estimate, 2 if a quarter's forecast is a range estimate, 1 if a quarter's forecast is an open-ended estimate, and 0 if a quarter's forecast is a qualitative statement; and
$INVMILLS_{it}$ = inverse Mills ratio based on an initial logit model estimated over initial decision of whether to give guidance in year *t*.

All other variables are defined in Table 3 and Table 4.

following the provision of guidance. The coefficients on other treatment variables (*ABSERROR_MG*, *CHGDISP*, and *MoB_AF*) are generally insignificant in both reduced and full models.[19] Overall, results from model (3) provide only weak support for a relation between forecast form and prior adverse outcomes. The next section investigates the extent to which these results are influenced by prior forecast form.

In summary, results are consistent with our predictions that adverse outcomes affect whether a firm guides at all next period and the number of fiscal quarters for which firms give guidance. Two patterns warrant further discussion. First, we find that managers' ability to meet or beat their own forecasts (MoB_MG_{it-1}) is significant in model (1) but insignificant in model (2). This suggests that firms falling short of their own forecasts are less likely to issue any guidance in the next year. However, conditional on forecasting at least once in the next fiscal year, the ability to meet or beat one's own forecasts is unrelated to the number of quarters for which firms issue guidance in that fiscal year. Second, we measure market disappointments at earnings announcements using both analyst forecasts (MoB_AF_{it-1}) and price reactions (CAR_EA_{it-1}). In models (1) through (3), one of these two is typically significant (or marginally significant), although the exact proxy for disappointment that is significant varies across the three models.

V. SPECIFICATION CHECKS AND ADDITIONAL EVIDENCE

Management Responses Conditioned on Prior Forecast Form

The evidence in Section IV suggests that prior adverse outcomes influence the frequency of future guidance, but provides limited support for a relation between prior adverse outcomes and future forecast form. This section investigates the extent to which these results vary with prior forecast form. The relation between adverse outcomes and prior forecast form is complicated by "ceiling effects" and "floor effects" related to prior forecast form. If firms issued imprecise forecasts in the prior year, then they have limited options to become less precise (floor effects). Similarly, if firms issue precise forecasts last year, then they have limited options to become more precise (ceiling effects). When firms with adverse outcomes cannot effectively respond by adjusting their forecast form (due to ceiling and floor effects), they may instead react to adverse outcomes by omitting future forecasts.

We investigate the effects of prior forecast form by partitioning our sample into firms with an average lagged form less than or equal to 2 (firms imprecise last year) and firms with an average lagged form greater than 2 (firms precise last year). Recall that forecast form is coded as 3, 2, 1, and 0 when the forecast is a point, range, open-ended, and qualitative forecast, respectively. A lagged average form greater than 2 therefore indicates that a firm gave at least one point estimate in the prior year and was, on average, more specific than a firm with a lagged average form equal to or less than 2.[20] We then re-estimate models (1) and (3) for each subsample, to investigate how guidance omission and form decisions are conditioned on prior forecast form.[21] When re-estimating model (3) we omit lagged form as an independent variable, as the subsamples were formed by partitioning on lagged form and thus lagged form has little variation within each subsample.

---

[19] When re-estimating model (3) using principal component analysis we find a significantly positive coefficient on the factor related to meeting or beating earnings benchmarks ($\beta_1 = 0.0375$, p = 0.04) and an insignificant coefficient on the factor related to absolute management forecast errors and a lack of analyst response ($\beta_2 = -0.0291$, p = 0.20).

[20] We consider the distinction between ranges and point estimates to be the key factor in selecting a cut-off for classifying prior forecast form, as point estimates cannot become more specific. In addition, the high frequency of point estimates and ranges (as compared to the relatively rare open-ended and qualitative forecasts) suggests that for most firms, the key decision regarding forecast form is whether to issue a point versus a range forecast.

[21] For this additional analysis, we omit model (2) regarding the number of quarters within a year for which a firm issues forecasts because we feel that predictions about how within-year frequency would vary based on prior forecast form are more ambiguous.

Results from re-estimating models (1) and (3) for each subsample are presented in Table 7. Re-estimating model (1) for imprecise prior forecasts ($FORM_{it-1} \leq 2$) yields results qualitatively similar to the full sample results. In particular, MoB_MG_{it-1} and CAR_EA_{it-1} remain negatively related to forecast omissions, while $DIFF\_AFMG_{it-1}$ and $CHGDISP_{it-1}$ remain positively related to forecast omissions. These results suggest that firms that were imprecise in the past respond to adverse outcomes by omitting future guidance. In contrast, when re-estimating model (3) for firms that were imprecise in the prior year, we find that MoB_MG_{it-1} (our only treatment variable with a directional prediction) is not significantly related to specificity. In other words, because these firms issue imprecise forecasts in prior years, they cannot further reduce the precision of future forecasts in an attempt to increase the likelihood of meeting or beating their own forecasts. Taken together, these results suggest that imprecise prior forecasters respond primarily to adverse outcomes by omitting forecasts, rather than further reducing forecast specificity.

When re-estimating model (1) for precise prior forecasts ($FORM_{it-1} > 2$), the coefficients on most treatment variables become insignificant. Therefore, when firms were precise in the prior year, past adverse outcomes are generally not significantly related to omissions. A notable exception is the positive relation between omissions and $DIFF\_AFMG_{it-1}$ ($\beta_3 = 68.8437$, p = 0.04). That is, when analysts do not closely follow precise forecasts, managers respond by omitting forecasts. This makes sense, as managers' ability to become more precise is bounded given that they have already been precise. If analysts do not follow precise guidance, then they seem unlikely to follow imprecise guidance, so managers respond by omitting guidance rather than reducing specificity. When re-estimating model (3) using firms that were precise last year, we find a marginally positive relation between MoB_MG_{it-1} and future forecast form. Therefore, firms that were precise in the past and fell short of their own guidance respond by being less precise in the future in order to increase the likelihood of meeting or beating future guidance.

In sum, results from re-estimating models (1) and (3) conditioning on prior forecast form suggest that management's reaction to adverse outcomes varies with prior specificity. If managers were imprecise in the past, adverse outcomes lead them to omit future guidance. If managers were precise in the past, they will reduce specificity if they failed to meet or beat their own guidance, or will omit guidance if analysts deviated from management's precise guidance.

Quarterly and Two-Year Rolling Windows

The results of Section IV are based on annual horizons that compare outcomes from year $t-1$ to guidance behavior in year t. We next examine the robustness of our full-model results to two alternate windows: quarterly analyses and two-year rolling windows. We first re-estimate models (1) and (3) using quarterly data in which outcomes from quarters $q-4$ to $q-1$ are used to predict guidance behavior in quarter q. The quarterly version of *OMIT* takes a value of 1 if the firm does not issue any quarterly guidance in quarter q, and 0 otherwise. The quarterly version of *FORM* is based on the specificity of guidance issued during quarter q. Variable definitions for all independent variables are similar to those given for models (1) and (3) except that variables previously calculated over year $t-1$ are instead calculated over quarters $q-4$ to $q-1$. Quarterly results are generally similar to the annual results for all treatment variables in model (1). When re-estimating model (3) using quarterly results, *MoB_MG* becomes significantly positive ($\beta_1 = 0.0729$, p = 0.05), consistent with the predictions of H1. Using annual data, this relation is only significant in the reduced model. *CAR_EA* becomes significantly positive ($\beta_6 = 0.1902$, $p < 0.01$), while this is only marginally significant using annual data.

We next re-estimate models (1) through (3) by aggregating prior outcomes and forecasting decisions over two-year horizons. For this analysis, we use a rolling window in which outcomes from years $t-2$ and $t-1$ are used to predict guidance behavior for years t and $t+1$. *OMIT* takes a value of 1 if the firm does not issue any quarterly guidance in years t and $t+1$ and 0 otherwise.

TABLE 7

Summary of Coefficient Estimates (Robust z-statistics) Conditioning on Prior Forecast Form

	Model (1): Forecast Omissions			Model (3): Forecast Specificity		
Variable	**Predicted Sign**	**$FORM_{it-1} \leq 2$**	**$FORM_{it-1} > 2$**	**Predicted Sign**	**$FORM_{it-1} \leq 2$**	**$FORM_{it-1} > 2$**
MoB_MG_{it-1}	−	−0.5596***	−0.0342	+	0.0634	0.2055*
		(−3.90)	(−0.12)		(1.23)	(1.76)
$ABSERROR\_MG_{it-1}$	+	−4.8215	20.6634		0.3437*	−0.2588
		(−0.61)	(1.38)		(1.71)	(−0.88)
$DIFF\_AFMG_{it-1}$	+	28.6439*	68.8437**		−18.5658***	−15.8620
		(1.93)	(2.04)		(−3.16)	(−0.97)
$CHGDISP_{it-1}$	+	50.6693**	14.8234		−5.2879	−21.8604
		(2.05)	(0.29)		(−0.87)	(−1.25)
MoB_AF_{it-1}	−	0.0410	−0.0433		−0.0178	0.0498
		(0.26)	(−0.14)		(−0.34)	(0.39)
CAR_EA_{it-1}	−	−0.9197***	−0.6924		0.0791	0.3107
		(−2.66)	(−1.20)		(0.75)	(1.36)
$STDRET_{it-1}$		19.4454***	22.0057***			
		(5.30)	(3.37)			
$DISP_{it-1}$		24.1634	−47.6781			
		(1.29)	(−1.33)			
RET_{it-1}		−0.2113*	−0.4936**			
		(−1.85)	(−2.11)			
$LOSS_{it-1}$		0.1950	0.7313**			
		(1.06)	(2.46)			
$INCREASE_{it-1}$		−0.2478*	0.3976*			
		(−1.78)	(1.83)			
$NUMAN_{it-1}$		−0.0502***	−0.0503***			
		(−3.91)	(−2.84)			

(continued on next page)

TABLE 7 (continued)

Variable	Model (1): Forecast Omissions			Model (3): Forecast Specificity		
	Predicted Sign	$FORM_{it-1} \leq 2$	$FORM_{it-1} > 2$	Predicted Sign	$FORM_{it-1} \leq 2$	$FORM_{it-1} > 2$
LIT_{it-1}		−0.3679***	−0.4304**			
		(−3.23)	(−2.57)			
$SIZE_{it-1}$		0.1193***	0.1056*			
		(3.14)	(1.94)			
BM_{it-1}		0.0410	−0.1205			
		(0.37)	(−0.67)			
$BETA_{it-1}$		−0.3169***	−0.6709***			
		(−3.20)	(−3.77)			
$DAYS_{it}$					0.0143	0.1165***
					(0.73)	(2.98)
$STDRET_{it-1}$					−0.2820	−7.1117***
					(−0.28)	(−3.09)
$CARSIGN_{it}$					0.0627*	0.0298
					(1.76)	(0.44)
$NUMAN_{it-1}$					0.0084***	0.0080
					(2.80)	(1.03)
$SIZE_{it-1}$					−0.0223**	−0.0872***
					(−2.07)	(−3.70)
$INVMILLS_{it}$					−0.0567	−0.0223
					(−0.43)	(−0.07)
Intercept		−1.3814***	−1.5074***		1.9998***	2.4332***
		(−4.18)	(−2.92)		(18.33)	(9.88)
# Observations		2,978	1,220		2,174	883
Model Chi-Square		217.13	96.31			
Adj. R^2					0.0315	0.1016

(continued on next page)

TABLE 7 (continued)

*, **, *** Indicates statistical significance at the 0.10, 0.05, and 0.01 levels, respectively, under two-tailed tests.
Robust z-statistics are based on standard error estimates that control for firm clustering effects.
All variables are defined in Table 3, Table 4, and Table 6.

NUMQTRS can take a value from 1 to 8, based on the number of quarters in years t and $t+1$ for which the firm issues guidance. *FORM* is based on the average of guidance specificity over each quarter of years t and $t+1$ for which the firm issues guidance. Variable definitions for all independent variables are similar to those given for models (1) through (3) except that variables previously calculated over year $t-1$ are instead calculated over $t-2$ and $t-1$.

When re-estimating model (1) using a two-year rolling window, the coefficients on *MoB_MG* and *CHGDISP* are insignificant ($\beta_1 = -0.1290$, p = 0.38; $\beta_4 = 30.2898$, p = 0.19), while these coefficients were significant in the annual results. The coefficient on *MoB_AF* is significantly negative in the two-year rolling window results ($\beta_5 = -0.4316$, p = 0.01), while this coefficient was insignificant in the annual results. Other two-year rolling window results are qualitatively similar to the annual results reported in Table 4. Overall, these results are consistent with a link between prior outcomes and future guidance omissions, although the exact factors that matter differ between two-year rolling windows and annual tests.

When re-estimating model (2) using a two-year rolling window, the coefficients on *MoB_MG* and *CAR_EA* are positive and significant ($\beta_1 = 0.1896$, p = 0.01; $\beta_6 = 0.4703$, p < 0.01), while these coefficients were insignificant in the annual results. All other two-year rolling window results for model (2) are qualitatively similar to the annual results. Overall, two-year rolling window results are consistent with a link between prior outcomes the number of quarters for which forecasting firms provide guidance. Full model results for re-estimating model (3) using two-year rolling windows suggest that our treatment variables are not significantly related to forecast form over this horizon.

Alternate Measures of Accuracy

One important outcome examined in this study is the extent to which managers are able to meet or beat their own forecasts. Our main analysis uses First Call's conversion of range, open-ended, and qualitative forecasts into point estimates to determine whether managers meet or beat their own forecasts. To check the robustness of our results to the treatment of imprecise guidance, we code range observations as meeting or beating management's forecast when actual earnings were greater than or equal to the lower bound of the range. Similarly, minimum forecasts are coded as meeting or beating management's forecast when actual earnings were greater than or equal to the minimum forecast.[22] Results using this alternate criterion for coding *MoB_MG* are qualitatively similar to those reported in Table 4.

As an additional sensitivity check, we create an alternate proxy for management's ability to meet or beat their own forecast, in which weight given to the outcome from a quarter is increasing in the precision of that forecast. We recalculate MoB_MG_{it-1} by multiplying $MoB\_MG_{it-1,q}$ for each quarter for which management issued earnings guidance in the previous year by 4 if that forecast was a point estimate, 3 for a range estimate, 2 for an open-ended estimate, or 1 for a qualitative forecast. The summed values of weighted $MoB\_MG_{it-1,q}$ are then scaled by the sum of the weights. Results using this alternate criterion for coding *MoB_MG* are qualitatively similar to those reported in Table 4. In summary, all results on the relation between guidance omissions and past outcomes are robust to alternate specifications of *MoB_MG*.

Seasoned Equity Offerings

Our final set of tests examines whether the sensitivity of forecasting omissions to prior outcomes varies with future seasoned equity offerings. Prior research (e.g., Frankel et al. 1995)

[22] Under this alternate criterion, maximum forecasts are always met or beaten, as they provide an upper bound on predicted future earnings. Our results are unaffected by the exclusion of maximum forecasts.

suggests that legal sanctions and other institutional penalties to optimistic forecasting are particularly severe when such forecasts are made shortly prior to equity offerings. As a result, we predict that managers with a history of falling short of their own guidance or disappointing market participants at earnings announcements are more likely to discontinue forecasting in the year before a seasoned equity offering.[23]

To test this conjecture, we examine whether the relation of $OMIT_{it}$ with MoB_MG_{it-1}, MoB_AF_{it-1}, and CAR_EA_{it-1} varies with whether a firm makes a seasoned equity offering in year $t+1$. This test presumes that managers are aware that a seasoned equity offering will occur in year $t+1$ when making guidance decisions in year t. Using the SDC seasoned equity offering database, we construct a new variable, SEO_{it+1} that takes a value of 1 if the firm i issues a seasoned equity offering in year $t+1$, and 0 otherwise. We re-estimate model (1), adding four additional independent variables: SEO_{it+1}, SEO_{it+1} interacted with MoB_MG_{it-1}, SEO_{it+1} interacted with MoB_AF_{it-1}, and SEO_{it+1} interacted with CAR_EA_{it-1}.

Results (untabulated) are similar to those in Table 4, as the coefficient on MoB_MG_{it-1} remains significantly negative, $ABSERROR\_MG_{it-1}$ remains insignificant, $DIFF\_AFMG_{it-1}$ remains positive and significant, $CHGDISP_{it-1}$ remains positive and significant, MoB_AF_{it-1} remains insignificant, and CAR_EA_{it-1} remains negative and significant. We find a marginally significant negative coefficient on the interaction between SEO_{it+1} and MoB_MG_{it-1} ($\beta = -0.9860$, p = 0.09), which suggests that the negative relation between forecast omissions and past ability to meet or beat one's own forecast is even more pronounced for firms that will have a seasoned equity offering in the following year. Coefficients on SEO_{it+1}, SEO_{it+1} interacted with MoB_AF_{it-1}, and SEO_{it+1} interacted with CAR_EA_{it-1} are all insignificant. In summary, these results are consistent with the prediction that legal liability concerns associated with optimistic forecasting are more pronounced prior to a seasoned equity offering. Legal liability concerns associated with market disappointments do not change prior to a seasoned equity offering.

VI. CONCLUSIONS

This study examines how outcomes of previous earnings forecasts affect a firm's quarterly forecasting strategy, including forecasting incidence, frequency, and specificity. Specifically, we examine four aspects of adverse outcomes. First, we argue that earnings guidance generates costs for the firm (in the form of legal liability, reputation loss, and stock price declines) when realized earnings fall short of management earnings predictions. Second, we argue that guidance is unlikely to be effective when analysts do not closely follow management's forecast. Third, we argue that managers are discouraged from future forecasting when guidance is unsuccessful at reducing information asymmetry. Fourth, we argue that firms face adverse consequences from issuing earnings guidance that results in subsequent disappointments at earnings announcements.

Our empirical work finds that firms are more likely to omit quarterly forecasts when facing each of the above adverse outcomes. The likelihood a firm omits quarterly forecasts for a given year is decreasing in the firm's past ability to meet or beat their own forecasts, increasing in the extent to which post-guidance analyst forecasts deviated from prior management guidance, increasing in the extent to which prior guidance increased the dispersion in analyst forecasts, and decreasing in the past ability to avoid earnings disappointments after issuing guidance. For firms that continue to give guidance, adverse prior outcomes also affect the precision of future guidance

[23] Alternatively, adverse outcomes could be less associated with forecast omissions prior to equity offerings if managers' primary concerns at this time are to correct prices and reduce information asymmetry. However, Frankel et al. (1995) find that firms are not more likely to issue forecasts in the period shortly before equity offerings. They interpret this as evidence that benefits to cost of capital from reducing information asymmetry through pre-offering forecasts are offset by the increased legal risk associated with such forecasts.

and the number of quarters within a year for which they give guidance. Prior research has suggested a number of firm-specific factors associated with decisions about providing guidance. Our study adds to existing literature by documenting that forecasting decisions are also influenced by the outcomes from previously issued earnings guidance.

REFERENCES

Ajinkya, B., and M. Gift. 1984. Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research* 22: 425–444.

———, S. Bhojraj, and P. Sengupta. 2005. The association between outside directors, institutional investors and the properties of management earnings forecast. *Journal of Accounting Research* 43: 343–376.

Anilowski, C., M. Feng, and D. Skinner. 2007. Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. *Journal of Accounting and Economics* 44: 36–63.

Baginski, S., E. Conrad, and J. Hassell. 1993. The effects of management forecast precision on equity pricing and on the assessment of earnings uncertainty. *The Accounting Review* 68: 913–927.

———, and J. Hassell. 1997. Determinants of management forecast precision. *The Accounting Review* 72: 303–312.

Bamber, L., and Y. Cheon. 1998. Discretionary management forecast disclosures: Antecedents and outcomes associated with forecast venue and forecast specificity choices. *Journal of Accounting Research* 36: 167–190.

Barron, O., and P. Stuerke. 1998. Dispersion in analysts' earnings forecasts as a measure of uncertainty. *Journal of Accounting, Auditing & Finance* 13: 243–268.

Brown, L., and M. Caylor. 2005. A temporal analysis of quarterly earnings thresholds: Propensities and valuation consequences. *The Accounting Review* 80: 423–440.

Chen, S., D. Matsumoto, and S. Rajgopal. 2007. Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance. Working paper, University of Texas at Austin.

Clement, M., R. Frankel, and J. Miller. 2003. Confirming management earnings forecasts, earnings uncertainty, and stock returns. *Journal of Accounting Research* 41: 653–679.

Coller, M., and T. Yohn. 1997. Management forecast and information asymmetry: An examination of bid-ask spreads. *Journal of Accounting Research* 35: 181–191.

Core, J. 2001. A review of the empirical disclosure literature: Discussion. *Journal of Accounting and Economics* 31: 441–456.

Cotter, J., I. Tuna, and P. Wysocki. 2006. Expectations management and beatable targets: How do analysts react to public earnings guidance? *Contemporary Accounting Research* 23: 593–628.

Diamond, D., and R. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *The Journal of Finance* 46: 1325–1359.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *The Journal of Finance* 59: 1553–1583.

Field, L., M. Lowry, and S. Shu. 2005. Does disclosure deter or trigger litigation? *Journal of Accounting and Economics* 3: 487–507.

Fields, T., T. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31: 255–307.

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32: 137–164.

Frankel, R., M. McNichols, and G. P. Wilson. 1995. Discretionary disclosure and external financing. *The Accounting Review* 70: 135–150.

Gebhardt, W., C. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39: 135–176.

Graham, J., C. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40: 3–73.

Han, B., and D. Manry. 2000. The implications of dispersion in analysts' earnings forecasts for future ROE and future returns. *Journal of Business Finance & Accounting* 27: 99–125.

Harris, M., and A. Raviv. 1993. Differences of opinion make a horse race. *Review of Financial Studies* 6: 473–506.

Hirst, E., L. Koonce, and S. Venkataraman. 2008. Management earnings forecasts: A review and framework. *Accounting Horizons* 22: 315–338.

Houston, J., B. Lev, and J. Tucker. 2010. To guide or not to guide? Causes and consequences of stopping quarterly earnings guidance. *Contemporary Accounting Research* 27: 143–185.

Hutton, A., and P. Stocken. 2009. Effect of reputation on the credibility of management forecasts. Working paper, Boston College.

———, and J. Weber. 2001. *Progressive Insurance: Disclosure Strategy*. Harvard Business School Case. Cambridge, MA: Harvard Business School Press.

Irani, A., and I. Karamanou. 2003. Regulation Fair Disclosure, analyst following, and analyst forecast dispersion. *Accounting Horizons* 17: 15–29.

Kandel, E., and M. Pearson. 1995. Differential interpretation of information and trade in speculative markets. *The Journal of Political Economy* 103: 831–872.

Kasznik, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 37: 57–81.

———, and B. Lev. 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70: 113–134.

King, R., G. Pownall, and G. Waymire. 1990. Expectations adjustment via timely management forecasts: Review, synthesis, and suggestions for future research. *Journal of Accounting Literature* 9: 113–144.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31: 105–231.

Lang, M., and R. Lundholm. 1993. Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research* 31: 246–271.

———, and ———. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71: 467–492.

Lennox, C., and C. Park. 2006. The market penalties for missing management forecasts and for missing guided analyst forecasts. Working paper, Nanyang Technological University.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77: 483–514.

Miller, G. 2002. Earnings performance and discretionary disclosure. *Journal of Accounting Research* 40: 173–204.

Nagar, V., D. Nanda, and P. Wysocki. 2003. Discretionary disclosure and stock-based incentives. *Journal of Accounting and Economics* 34: 283–309.

Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22: 435–480.

Richardson, S., S. Teoh, and P. Wysocki. 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21: 885–924.

Rogers, J., and P. Stocken. 2005. Credibility of management forecasts. *The Accounting Review* 80: 1233–1260.

Sharpe, W. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19: 425–442.

Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32: 38–60.

———. 1997. Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics* 23: 249–282.

———, and R. Sloan. 2002. Earnings surprises, growth expectations, and stock returns, or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7: 289–312.

Soffer, L., R. Thiagarajan, and B. Walther. 2000. Earnings preannouncement strategies. *Review of Accounting Studies* 5: 5–26.

Trueman, B. 1986. Why do managers voluntarily release earnings forecasts? *Journal of Accounting and Economics* 8: 53–71.

Verrecchia, R. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5: 179–194.

———. 2001. Essays on disclosure. *Journal of Accounting and Economics* 32: 97–180.

Waymire, G. 1986. Additional evidence on the accuracy of analyst forecasts before and after voluntary management earnings forecasts. *The Accounting Review* 61: 129–142.

Williams, P. 1996. The relation between a prior earnings forecast by management and analyst response to a current management forecast. *The Accounting Review* 71: 103–113.

Wooldridge, J. 2001. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 1985–2009

American Accounting Association
DOI: 10.2308/accr.00000002

Analyst Information Acquisition and Communication

Paul E. Fischer
The Pennsylvania State University

Phillip C. Stocken
Dartmouth College

ABSTRACT: We examine a communication game between an analyst and a decision-maker and investigate how the presence of public information affects the precision of the information the analyst gathers and communicates to the decision-maker. We characterize conditions under which public information causes the analyst to underinvest or overinvest in the information gathered relative to the case where analyst credibility is not an issue. We then discuss when the presence of public information causes the analyst to reduce the depth of coverage of the firm, suggesting that the introduction of public information can make the decision-maker strictly worse off.

Keywords: *information acquisition; cheap-talk; analyst credibility.*

I. INTRODUCTION

Decision-makers obtain information from various sources. Some sources have incentives to gather information and communicate in a self-serving fashion and face limited regulation regarding these activities. Other sources have little incentive to behave opportunistically or face fairly stringent regulation. Investors, for instance, can obtain information from sell-side analysts, who are thought to have incentives to communicate strategically and face relatively little regulation regarding their reporting behavior, and also from a firm's audited financial statements, where regulation plays a significant role in limiting strategic behavior. This study analyzes how changes in the information disclosed by a non-strategic source influence the information-gathering and communication decisions of a strategic information intermediary.

The model we employ has four stages and two players: an information intermediary, called the analyst, and a decision-maker. In the first stage, both players observe an imperfect public signal about the state of nature. In the second stage, the analyst engages in costly information gathering that determines the precision of a signal about the state of nature he will privately observe. In the third stage, the analyst privately observes a signal and costlessly sends a message

We thank Tim Baldenius, Ron Dye, Ilan Guttman, Steve Huddart, Nahum Melumad, John Morgan, Ram Ramanan, Jonathan Rogers, Richard Sansing, Jerry Salamon, Ulf Schiller, Reed Smith, and Ro Verrecchia for helpful discussion. We also thank workshop participants at the University of Bern, Chicago-Minnesota Accounting Theory Conference, Columbia University, Dartmouth College, Indiana University, University of Michigan, and University of Utah Winter Conference.

Editor's note: Accepted by Madhav Rajan.

Submitted: November 2005
Accepted: April 2010
Published Online: November 2010

to the decision-maker. In the fourth stage, the decision-maker, given her beliefs about the state of nature, takes an action affecting both players' payoffs. The players have divergent preferences regarding the decision-maker's action. Within the context of this model, we explore how the precision of the public information the players observe in the first stage and how the public observability of the analyst's precision choice in the second stage affects the analyst's information gathering and communication.

While the model can be applied to various settings in which decision-makers obtain information from both strategic and non-strategic sources, we find this model useful for framing the recent evolution in the sell-side equity analyst market. The amount of information firms make available directly to investors has exploded recently. Following the introduction of Reg FD in 2000, for instance, firms can no longer privately reveal information to analysts and, hence, are more likely to publicly release their forward-looking information (see Heflin et al. 2003). Further, firms now routinely hold open conference calls following their earnings releases (Bushee et al. 2003). In addition, the $1.4 billion settlement between the large brokerage firms and the State of New York in 2002 required the firms to provide information to investors from unbiased sources together with their own analysis. Following this settlement, there has been a flight of equity analysts from the research departments at the large investment firms, and entire industry sectors have experienced a reduction in analyst coverage. As a consequence, several commentators have argued that these changes are likely to hurt individual investors and reduce stock price efficiency (see Schack 2003). A contribution of this study is to illustrate how alternative sources of information might affect analyst information-gathering activities, the integrity of their communication with investors, and whether investors are likely to be better off as a consequence of changes in the amount of publicly available information.

Our analysis suggests that improvements in the quality of public information have subtle consequences for analyst behavior. When investors know the precision of the analyst's information (but not the signal realization), we find that changes in publicly observed information affect the analyst's behavior as follows. On one hand, when credibility issues are minimal, an increase in the precision of public information causes it to substitute for the analyst's information. As information is costly to gather, the analyst reduces the precision of his information. Although more precise public information crowds out some analyst information, this crowding out is insufficient to offset the benefit of more precise public information. Consequently, the decision-maker is better off. On the other hand, when the analyst's credibility is in doubt—because the incentives of the analyst and decision-maker are misaligned—the analyst's behavior depends on the cost of obtaining information that is sufficiently precise for it to be credibly communicated. Specifically, when information-gathering costs are *moderate*, more precise public information leads the analyst to gather more precise information to increase the decision-maker's responsiveness to his report and thereby facilitate credible communication. The heightened precision of the analyst information coupled with more precise public information combine to make the decision-maker better off. Alternatively, when information-gathering costs are *high*, it becomes exorbitantly costly for the analyst to gather sufficiently precise information to be able to credibly communicate. In this case, an increase in the precision of public information completely crowds out the analyst's information. Further, even though the decision-maker obtains more precise public information, the crowding out of the analyst's information can be so severe that it reduces the total information the decision-maker obtains and leaves her worse off.

When the analyst privately chooses the precision of information to gather, the analyst's information-gathering behavior depends on the realization of the public signal. Specifically, after observing favorable public information, the analyst gathers and credibly communicates more

precise information when the precision choice is not publicly observable than when it is publicly observable. Alternatively, after observing unfavorable public information, the analyst gathers and credibly communicates less precise information.

The primary theoretical antecedent of our analysis is Crawford and Sobel (1982). They consider a communication game between an informed expert and an uninformed decision-maker. The expert perfectly observes the realization of a payoff-relevant state variable and sends a costless message to the decision-maker, who then takes an action that affects both players. Their cheap-talk model has been extended in various directions. In work most closely related to ours, Austen-Smith (1994) considers a communication game in which the expert chooses whether to perfectly observe the state of nature at some fixed cost and decides whether to reveal being informed, but is unable to prove being uninformed. He establishes that there exist circumstances under which the players cannot communicate when the receiver knows the expert is informed, although they can do so when the receiver is uncertain about whether the expert is informed. In contrast to Austen-Smith (1994), we consider a model in which the expert, after observing a public signal, can choose the precision of his privately observed signal; further, we allow for the possibility that the precision of the expert's signal is either publicly observed or privately chosen.

The extant literature also has extended the model in Crawford and Sobel (1982) by examining the effect of multiple experts on the communication game (e.g., Austen-Smith 1993; Krishna and Morgan 2001; Wolinsky 2002; Battaglini 2004; Morgan and Stocken 2008), by introducing multiple decision-makers (e.g., Farrell and Gibbons 1989; Newman and Sansing 1993; Gigler 1994), or by allowing the players to exchange messages (e.g., Aumann and Hart 2003; Krishna and Morgan 2004). In these studies, the expert is exogenously endowed with information. In contrast, we examine a setting where it is costly for the expert to gather information and consider how the presence of a non-strategic information source affects the information the expert gathers and communicates.

Our analysis is also related to the voluntary disclosure literature. Verrecchia (1990) examines how a manager's discretionary disclosure behavior is influenced by the exogenous quality of his privately observed information. In contrast, Penno (1996), Pae (1999), and Hughes and Pae (2004) extend this model to consider a setting in which the manager makes a costly information-acquisition decision, which endogenizes the quality of the manager's information. A common feature of these studies, like much of the extant literature examining voluntary disclosure behavior (e.g., Verrecchia 1983; Dye 1985; Arya and Mittendorf 2007; Einhorn and Ziv 2008), is that private information, if it is disclosed, must be disclosed truthfully. In the analyst reporting environment, however, analysts often report in a self-serving fashion and face little regulation regarding how they report (e.g., Michaely and Womack 1999; Lin and McNichols 1998). Indeed, this concern motivated key provisions of the Sarbanes-Oxley Act as well as the legal settlement between the large brokerage firms and the State of New York in 2002. We examine a reporting environment in which the analyst's disclosure is not constrained by particular disclosure rules and find the desire to communicate credibly fundamentally influences the analyst's information-gathering behavior. Hence, a contribution of our analysis is to provide insights into information-gathering and communication behavior when reporting credibility is a key feature of the environment.[1]

Section II describes the model. Section III characterizes the analyst's precision choice when this choice is observable and the analyst is constrained to report truthfully. Section IV continues to

[1] Some extant work has studied analyst information gathering. Hayes (1998), for instance, examines an analyst's incentives to gather and report information in an environment in which the analyst's report influences an investor's trading behavior that, in turn, determines the trading commission that the analyst receives. The analyst always reports truthfully so that strategic disclosure, which is the focus of this paper, does not arise.

examine a setting in which the analyst's precision choice is observable, but relaxes the assumption that the analyst must report truthfully. It examines how changes in the precision of public information influence an analyst's information gathering and communication, and how these changes affect the quality of the decision-maker's information. Section V further extends the model to consider a setting which the decision-maker does not observe the precision of the analyst's information. It contrasts the precision of information the analyst gathers in this case relative to the case in which the precision of the analyst's information is publicly observable. Section VI concludes. All proofs are in the Appendix.

II. MODEL

Consider a communication game in which an analyst chooses to obtain private information about an unknown state variable and sends a report to a decision-maker who then takes an action. The unknown state variable is represented by the random variable $\tilde{\alpha}$, which has support $\{s,f\}$ where $s > f$. The common prior beliefs are that the successful, s, and failure, f, state realizations are equally likely to occur.[2]

The game has four stages. In the first stage, information represented by the realization x of a random variable $\tilde{x}$ is publicly observed. The support for $\tilde{x}$ is $X = \{h,l\}$. The probability that $\tilde{x} = h$ conditional on $\tilde{\alpha} = s$ and the probability that $\tilde{x} = l$ conditional on $\tilde{\alpha} = f$ are both $q \in (1/2,1)$. The variable q captures the precision of $\tilde{x}$, where a higher value for q implies that $\tilde{x}$ is more informative. Setting $q = 1/2$ is equivalent to supposing that the public information is absent. The public information can be viewed as the filing of a firm's audited financial statements or the release of governmental statistics relevant to a firm's future outlook. The analyst observes the realization of the public information x before choosing the precision of his private information.

In the second or *information-gathering* stage, the analyst chooses how much to spend gathering information. The analyst's information, which he privately observes, is represented by the realization y of a random variable $\tilde{y}$, which has support $Y = \{g,b\}$. Conditional on the realization of $\tilde{\alpha}$, $\tilde{x}$ and $\tilde{y}$ are independent. The probability that $\tilde{y} = g$ conditional on $\tilde{\alpha} = s$ and the probability that $\tilde{y} = b$ conditional on $\tilde{\alpha} = f$ are both $p \in [1/2,1]$. The amount spent gathering information $c(p)$ is a function of the precision p of the information, where $c(p)$ is a twice differentiable function that satisfies $c'(p) > 0$ and $c''(p) > 0$ for $p > 1/2$; that is, the analyst's cost of gathering information increases in the precision of the information and at an increasing rate. The analyst choosing not to gather private information is analogous to the case in which the analyst's information has precision $p = 1/2$.[3]

In the third or *reporting* stage, the analyst costlessly sends a report r to the decision-maker who then takes an action $a \in (-\infty, \infty)$. Without loss of generality, we restrict the analyst's possible reports to the set $R = \{g,b\}$. The analyst's reporting strategy is a function over the public signal x and the analyst's information y such that $X \times Y \rightarrow R$. The decision-maker's strategy is a function over the publicly observed signal and the analyst's report such that $X \times R \rightarrow A$.

In the fourth stage, the state of nature α is realized and the analyst's and decision-maker's payoffs are determined. The decision-maker's payoffs are:

$$U_d = -(\alpha - a)^2. \tag{1}$$

Given this payoff function, it is optimal in the third stage for the decision-maker to choose an action $a = E[\tilde{\alpha} \mid x, r]$, where $E[\,.\mid.\,]$ is the expectation operator, which is analogous to investors

[2] As in this paper, many cheap-talk models assume the players' prior beliefs are diffuse (e.g., Austen-Smith 1993; Krishna and Morgan 2001; Battaglini 2004).

[3] In the context of this binary model, the precision p of the analyst's information affects both the mean and variance of $\tilde{\alpha}$ conditional on the realization of $\tilde{y}$.

valuing a firm at its expected value.[4] The analyst's payoffs are:

$$U_i = \phi a - (\alpha - a)^2 - c(p), \tag{2}$$

where $\phi > 0$.[5] When $\phi = 0$, the players' interests are perfectly aligned because, given the same information, both players prefer the same action in the third stage. When $\phi > 0$, however, their interests are misaligned because the analyst's preferred action exceeds the decision-maker's preferred action. These utility representations are broadly descriptive of environments in which an analyst wants to induce a higher action than a decision-maker prefers, but is constrained from inducing an action that is too high because of, say, reputation or litigation concerns (e.g., Dugar and Nathan 1995; Grossman and Helpman 2001; Morgan and Stocken 2003). For instance, sell-side equity analysts often have incentives to offer optimistic recommendations on target firms to curry favor with the firms' management. They are constrained from being too exuberant, however, because their employers often consider their ranking in the annual *Institutional Investor* All-American Research Team poll when setting their compensation. This poll of money managers and institutions is widely viewed as a measure of an analyst's reputation (Michaely and Womack 1999).[6]

Furthermore, consistent with the institutional observation that analysts do not update their stock reports in response to every public information release affecting a firm, the specification of the analyst's objective function implies that, even when the analyst does not collect any private information, the analyst's payoff continues to be affected by the decision-maker's actions. Hence, within the context of the analyst reporting environment, p reflects the depth of coverage.

All aspects of the game are common knowledge except the analyst's private signal y and, depending on the case, the analyst's precision choice. Figure 1 summarizes the time line of events and the model's notation.

We study *Perfect Bayesian Equilibria*, which require that players' beliefs satisfy Bayes' rule

FIGURE 1
Time Line of Events



Information gathering stage *Reporting stage*

The unknown state α assumes the values s or f with equal probability. In stage one, the analyst and decision-maker observe the public information x, which is either h or l, with probability $Pr(h|s) = Pr(l|f) = q$.

In stage two, the analyst chooses the precision p of his private information y, which is either g or b, with probability $Pr(g|s) = Pr(b|f) = p$. The cost function $c(p)$ reflects the analyst's cost of gathering information with precision p.

In stage three, the analyst observes his private information y and sends a report r to the decision-maker. The decision-maker then chooses an action a. The parameter φ denotes the extent to which the players' interests are misaligned.

In stage four, the state of nature α is realized, and the analyst's payoff U_i and the decision-maker's payoff U_d are determined.

[4] The presumption that investors value a firm at its expected value is standard in the voluntary disclosure literature (e.g., Verrecchia 1990; Penno 1996; Einhorn 2007).

[5] We can also assume $\phi < 0$ without altering the flavor of the results.

[6] Several recent regulatory actions were aimed at eliminating the reporting conflict between analysts and investors. The Sarbanes-Oxley Act, enacted on July 25, 2002, was drafted to strengthen the independence of security analysts (Razaee 2007). The legal settlement between the large brokerage firms and the State of New York in 2002 has separated research and investment banking activities within investment firms, altered the way in which analysts are compensated, and mandated disclosure in stock reports of conflicts of interest that might exist. Despite these actions, analysts continue to have incentives to bias their reports to curry favor with management. Mayew (2008) finds that management favors those analysts in conference calls who have bullish stock recommendations on the firm.

whenever possible and their actions maximize their expected payoffs. The decision-maker's action a maximizes her expected payoff $U_d = -E[(\tilde{\alpha} - a)^2 \mid x, r]$. The analyst's choice of information precision p maximizes $U_i = E[\phi a - (\tilde{\alpha} - a)^2 - c(p) \mid x]$, and when the cost of gathering information is sunk, the analyst's choice of report r maximizes $U_i = E[\phi a - (\tilde{\alpha} - a)^2 \mid x, y]$. It can be established that there do not exist equilibria in mixed strategies in which some communication occurs in the reporting stage. Hence, the exposition disregards the possibility of mixed strategies.[7]

III. BENCHMARK WITH TRUTHFUL REPORTING

We consider a benchmark setting in which the analyst must report truthfully, $r = y$, and his precision choice p is common knowledge. In the reporting stage, the decision-maker's action is $a = E[\tilde{\alpha} \mid x, r = y]$. Given this action, the analyst's precision choice maximizes:

$$E[E[\phi E[\tilde{\alpha}|x,y] - (\tilde{\alpha} - E[\tilde{\alpha}|x,y])^2 - c(p)|x,y]|x] = \phi E[\alpha|x] - \frac{p(1-p)(s-f)\Delta(p)}{2p-1} - c(p), \quad (3)$$

where $\Delta(p)$ is the difference between the decision-maker's actions when she believes she knows y, that is:

$$\Delta(p) \equiv E[\tilde{\alpha}|x,g] - E[\tilde{\alpha}|x,b] = \frac{q(1-q)(2p-1)(s-f)}{(p(1-q)+(1-p)q)((1-p)(1-q)+pq)}. \quad (4)$$

The difference $\Delta(p)$ is increasing in p, which implies that the decision-maker's action choice is more responsive to the analyst's information when the analyst's information is more precise. It is decreasing in q, which implies the decision-maker's action is less responsive to the analyst's information when the public information is more precise. The difference is independent of the realization of x, which stems from the assumption that the prior beliefs regarding the unknown state variable α are diffuse.[8]

To ensure a unique, interior optimal choice of precision, we assume the analyst's cost function $c(p)$ satisfies the following additional conditions for the remainder of the analysis: (i) $\lim_{p \to 1/2} \frac{c'(p)}{(p-1/2)} \to 0$; (ii) $\lim_{p \to 1} c'(p) \to \infty$; (iii) $\frac{c''(p)}{c'(p)} > \frac{1-3p(1-p)}{p(1-p)(p-1/2)}$ for all $p \in (1/2, 1)$. Condition (i) implies that when p is close to one-half, the marginal benefit of increasing the precision of the analyst's information is strictly greater than its marginal cost. Condition (ii) implies that when p is close to one, the marginal cost of increasing the precision of information exceeds the marginal benefit. Condition (iii) implies the analyst's objective function is quasi-concave (i.e., single-peaked), which ensures the analyst's precision choice is unique. The assumption requires that the analyst's cost function for gathering information must be sufficiently convex that a departure from the optimal choice of precision would cause the analyst's expected utility to strictly decline. The power function $c(p) = ((p - 1/2)/(1-p))^n$, where $n > 2$, is an example of a cost function that satisfies these conditions.

Given the precision choice p must be in the interior, that is $p \in (1/2,1)$, the analyst's optimal precision choice satisfies the first-order condition characterized in Lemma 1.

Lemma 1: When the analyst's precision choice is observable and the analyst must report truthfully, the analyst's unique precision choice p^* satisfies:

$$\Delta^2(p^*)/(2p^* - 1) - c'(p^*) = 0. \quad (5)$$

The analyst's choice of precision for the signal y is independent of the realization of the public

[7] Further details available from the authors on request.

[8] Details about the consequence of varying the prior beliefs are available from the authors on request.

information x. To develop intuition for this finding, notice that the analyst's objective function prior to observing the signal y, as given in (3), is increasing in the decision-maker's expected action and decreasing in the cost of inducing the decision-maker to take an inappropriate action. Consider the first term, $E[\alpha|x]$, in (3). The expected action is increasing in the realization of the public information x. However, a change in the precision of the anticipated signal y does not affect the decision-maker's expected action—it remains $E[\alpha|x]$. Consider the second term, $E[E[(\tilde{\alpha} - E[\tilde{\alpha} \mid x, y])^2 \mid x, y] \mid x]$, in (3). Here a change in precision of the anticipated signal y does affect the decision-maker's expected responsiveness to the analyst's message. However, the expected variance term does not vary in the realization of x because, conditional on the realization of $\tilde{\alpha}$, $\tilde{x}$, and $\tilde{y}$ are independent. It follows, therefore, that the analyst's optimal precision choice does not depend on the public information.

Analyst Behavior and Decision-Maker Information

The focus of our analysis is on the impact of the quality of the public information, q, on the analyst's behavior. Accordingly, we consider the relation between the analyst's precision choice, p^*, and the quality of the public information, q

Corollary 2: When the analyst's precision choice is observable and the analyst must report truthfully, the chosen precision of the analyst's information, p^*, is decreasing in the precision of the public information q; that is, $\partial p^*/\partial q < 0$.

The intuition for the inverse relation between q and p^* is familiar. As the precision of public information increases, it substitutes for the analyst's information, causing the marginal benefit of the analyst's information to fall. Since gathering information is costly, the choice p^* that maximizes the analyst's expected utility decreases.

The intuitive crowding-out finding in Corollary 2 suggests that the impact of any increase in public information on decision-maker information is tempered by a loss of analyst information. Given the countervailing decline in the analyst information associated with an increase in the public information, it is unclear whether an increase in public information will ultimately lead to an increase in decision-maker information. To formally assess the effect of the change in the public information on the decision-maker's information, however, requires a measure of the decision-maker's information. We define the *quality of information* as the decision-maker's expected posterior variance after observing the public signal and analyst report, $-E[var[\tilde{\alpha} \mid x, r] \mid x]$. The statistical measure, $-E[var[\tilde{\alpha} \mid x, y; p^*] \mid x]$, which is consistent with Crawford and Sobel (1982), is appealing because it is equivalent to the decision-maker's expected utility. The decision-maker's expected posterior variance after the public signal is realized and analyst report is observed is identical to the decision-maker's expected utility in the information-gathering stage after the public signal is realized, but before the analyst's information is observed.[9]

The decision-maker's quality of information when the analyst chooses $p = p^*$ is given by $-E[var[\tilde{\alpha} \mid x, y; p^*] \mid x] = -p^*(1-p^*)(s-f)\Delta(p^*)/(2p^*-1)$. Using this measure, we find an increase in the precision of the public information q leads to an increase in the decision-maker's information; that is:

Corollary 3: Suppose the analyst's precision choice is observable and the analyst must report truthfully. The decision-maker's quality of information is increasing in the precision of the public information, q.

[9] The decision-maker's utility is identical to the middle term in the analyst's objective function in Equation (3). Accordingly, this term serves to align, at least partially, the interests of the analyst and decision-maker.

IV. UNCONSTRAINED REPORTING

Analysts are fairly unconstrained in how they report their private information. Consistent with this observation, we alter the model and allow the analyst to report without cost either g or b, regardless of the realization of his private information y. We continue to assume that the analyst's choice of p is common knowledge to focus solely on how communication incentives affect analyst information-gathering decisions. This setting is arguably consistent with the sell-side equity analyst environment. Analyst stock reports vary in content and often contain detailed analysis of a company, its key competitors, and its industry. The detail in the analyst's stock report, which the investor can readily observe, reflects the precision of the analyst's information p. Investors, however, do not observe the analyst's procedures for evaluating and interpreting the information that is gathered. It is this interpretation and evaluation that provides the analyst with private information y. Later, we shall relax the assumption that p is common knowledge.

Like most cheap-talk games, there are multiple equilibria. There always exist equilibria in which the analyst does not gather any information. There may or may not also exist equilibria in which the analyst gathers and communicates information. When there exist equilibria in which information is gathered, we focus on this class because these equilibria *ex ante* Pareto dominate any equilibrium in which information is not gathered.[10] Within the class of equilibria in which information is gathered, we focus on the *truthful communication* equilibrium—the equilibrium in which $r = y$ for all y. This focus is without loss of generality because when there exists an equilibrium with full revelation that involves the analyst reporting something other than y (e.g., $r \in \{b,g\}$ and $r \neq y$ for all y), it is economically equivalent to the truthful communication equilibrium.

When the players' interests are perfectly aligned, communication is credible and the analyst can always fully reveal his privately observed information. When the players' interests are misaligned, however, an analyst cannot always fully reveal his privately observed information. To assess whether the analyst can credibly communicate y, note that the analyst prefers a higher action than the decision-maker, given y, and that the decision-maker would take a higher action if she believes the analyst has observed $\tilde{y} = g$. It follows that the analyst is always willing to reveal when $\tilde{y} = g$, but may want to mislead the decision-maker when $\tilde{y} = b$ by nefariously claiming that $\tilde{y} = g$. Recalling that the value of the analyst's objective is increasing and then decreasing in the decision-maker's action implies the analyst will not try to dissemble when $\tilde{y} = b$ if the action induced from claiming that $\tilde{y} = g$ is much higher than the action preferred by the analyst when $\tilde{y} = b$. Hence, the analyst can communicate y if the action he would induce by misrepresenting his information is sufficiently high relative to the one he induces if he truthfully communicates his information. Formally, the analyst can credibly communicate if and only if:

$$\Delta(p) \equiv E[\tilde{\alpha}|x,g] - E[\tilde{\alpha}|x,b] = \frac{q(1-q)(2p-1)(s-f)}{(p(1-q)+(1-p)q)((1-p)(1-q)+pq)} \geq \phi. \tag{6}$$

Observe that $\Delta(p)$ is an increasing function of p, which implies that the decision-maker's action choice is more responsive to the analyst's information when the analyst's information is more precise. As discussed above, the analyst can communicate y if the action he induces when he claims to have observed realization g is sufficiently large relative to the action he induces when he claims to have observed realization b. Hence, if p is sufficiently high, the analyst can communicate y. The observation that the analyst can fully reveal his privately observed information provided the

[10] It is standard in the cheap-talk literature to focus on an *ex ante* Pareto dominant equilibrium; see, for instance, Crawford and Sobel (1982) and Alonso et al. (2008).

players' interests are not too misaligned and the analyst's information is sufficiently precise is formalized in the next lemma.[11]

Lemma 4: Suppose the analyst's precision choice is observable and the analyst need not report truthfully. The analyst can communicate y in the presence of the other information x if and only if the precision of the analyst's information is sufficiently high; that is, there exists a minimum threshold $\bar{p}$ such that, for any x, the analyst can credibly communicate y if and only if $p \geq \bar{p}$.

When the analyst must truthfully report y, Lemma 1 established that the analyst's precision choice is maximized at p^*. When the analyst is not restricted to report truthfully, Lemma 4 implies that p must be at least $\bar{p}$ for credible communication to occur. It follows from these two lemmas that, when the analyst is not restricted to report truthfully, the analyst chooses information quality p^* if credible communication is possible, specifically $p = p^* \geq \bar{p}$. Otherwise, the analyst chooses a higher level of information quality to allow communication, $p = \bar{p} > p^*$, or he abstains from gathering information if information of higher quality is too costly, that is, $p = 1/2$. To formalize when the analyst decides to forgo gathering information as opposed to gathering information that is sufficiently precise to facilitate communication, we characterize a threshold $p^{\#}$. This threshold is the level of precision such that the analyst is indifferent between gathering information with precision $p^{\#}$ that is truthfully revealed and not gathering information, $p = 1/2$; formally, $p^{\#}$ solves:

$$c(p^{\#}) = \Delta(p^{\#})q(1-q)(2p^{\#}-1)(s-f). \tag{7}$$

Note that the threshold $p^{\#}$ strictly exceeds the interior choice p^*.

We characterize the analyst's equilibrium information-gathering and communication decisions as a function of p^*, $\bar{p}$, and $p^{\#}$ in Proposition 5.

Proposition 5: Suppose the analyst's precision choice is observable and the analyst need not report truthfully. There always exists an equilibrium in which the analyst does not gather any information, $p = 1/2$. Furthermore:

(i) If $\bar{p} \leq p^* < p^{\#}$, then there exists a unique equilibrium with an interior level of precision, which equals p^*. The analyst truthfully communicates the information gathered.

(ii) If $p^* < \bar{p} \leq p^{\#}$, then there exists a unique equilibrium with an interior level of precision, which equals $\bar{p}$. The analyst truthfully communicates the information gathered.

(iii) If $p^* < p^{\#} < \bar{p}$, then there does not exist an equilibrium with an interior level of precision.

A key feature of Proposition 5 is that the analyst's credibility problem alters the structure of the costs the analyst faces. When the analyst is unable to communicate, he cannot coordinate the decision-maker's actions. This inability is costly. By gathering more precise information, however, the analyst can credibly communicate information that causes the decision-maker to take more informed actions.[12] Since the players' interests are partially aligned, this communication not only

[11] The claim that the analyst can credibly communicate when p is sufficiently high can be established for a more general specification of the analyst's objective function $U(a,\alpha,\phi)$ that has the properties: (i) $\partial U(a,\alpha,\phi)/\partial a = 0$ for some action a; (ii) $\partial^2 U(a,\alpha,\phi)/\partial a^2 < 0$ so that the analyst's utility function $U(.)$ has a unique maximum action a for each pair (α,ϕ); and (iii) $\partial^2 U(a,\alpha,\phi)/\partial a \partial \alpha > 0$ so that the best decision-maker's action from the analyst's perspective is increasing in the state α. The more general utility function does not lend itself to comparative static exercises in a straight-forward manner, motivating the tractable utility function specified in Section II.

[12] In contrast, if the players' interests are sufficiently misaligned that $\phi > (s - f)$, then the analyst is unable to gather sufficiently precise information to credibly communicate. Hence, only the babbling equilibrium prevails. In this equilibrium, the analyst does not collect any private information.

heightens the decision-maker's expected utility, but also raises the analyst's expected utility. Consequently, the analyst's desire to influence the decision-maker motivates the analyst to gather information.

Analyst Behavior

When the analyst is constrained to report truthfully, increasing the public information, q, causes the analyst to reduce his information gathering. In contrast, when the analyst is not constrained to report truthfully, the effect of public information on the analyst's information-gathering activities is more subtle. Examining how the public information, q, influences the analyst's information-gathering behavior entails determining how q affects the critical values p^*, $\bar{p}$, and $p^{\#}$. The inverse relation between p^* and q was established earlier.

For the analyst to credibly communicate, he must gather information with precision of at least $\bar{p}$. Corollary 6 shows $\bar{p}$ is increasing in q.

Corollary 6: The minimum precision of the analyst's private information necessary for the analyst to credibly communicate, $\bar{p}$, is *increasing* in the precision of public information q; that is, $\partial \bar{p} / \partial q > 0$.

An increase in the precision of the public information must be accompanied by an increase in the precision of the analyst's private information for the decision-maker to remain responsive to the analyst's report and thereby encourage the analyst to continue to report truthfully. Conversely, an increase in the precision of the public information without a corresponding increase in the precision of the analyst's report causes the decision-maker to place relatively more weight on the public information and relatively less weight on the analyst's report. The analyst is then less capable of influencing the decision-maker's action and, accordingly, is more inclined to lie.

The analyst might decide not to gather information that is sufficiently precise to credibly communicate because it is too costly to do so. This decision hinges on the level of precision $p^{\#}$—the unique level of precision for which the analyst is indifferent between gathering information that can be communicated and not gathering information. This level of precision $p^{\#}$ also depends on the precision of the public information, q. As the precision of public information increases, the benefit to the analyst of providing the decision-maker with more precise information decreases. Hence, $p^{\#}$ declines. Corollary 7 formalizes this discussion.

Corollary 7: The level of precision of the analyst's private information for which the analyst is indifferent between gathering and not gathering information, $p^{\#}$, is *decreasing* in the precision of public information q; that is, $\partial p^{\#} / \partial q < 0$.

Having established the relations between precision of the public information, q, and analyst information precision levels p^*, $\bar{p}$, and $p^{\#}$, we now assess how the analyst's equilibrium precision choice varies in the precision of the public information. Consider the case for which the precision of the public information is relatively low, so that $p^* \geq \bar{p}$. As q increases, $\bar{p}$ rises and p^* falls. Because the analyst chooses p^* when $p^* \geq \bar{p}$, the analyst's choice of p falls in q because p^* falls in q. Eventually, q rises to a level where $p^* = \bar{p}$. As q continues to rise, communication is no longer credible if the analyst chooses p^*. Hence, the analyst must choose between the level of precision necessary for credible communication $\bar{p}$ or not collect any information by choosing $p = 1/2$. Initially, the analyst responds by choosing $\bar{p}$ and *overinvesting* in the quality of information he gathers to allow credible communication. Given that $\bar{p}$ is increasing in q, it follows that the precision of the analyst information increases in q. When $\bar{p} \geq p^{\#}$, however, implementing $\bar{p}$ becomes so costly that the analyst decides not to collect any information. In these circumstances, the improvement in the public information causes the analyst to *underinvest* in the quality of infor-

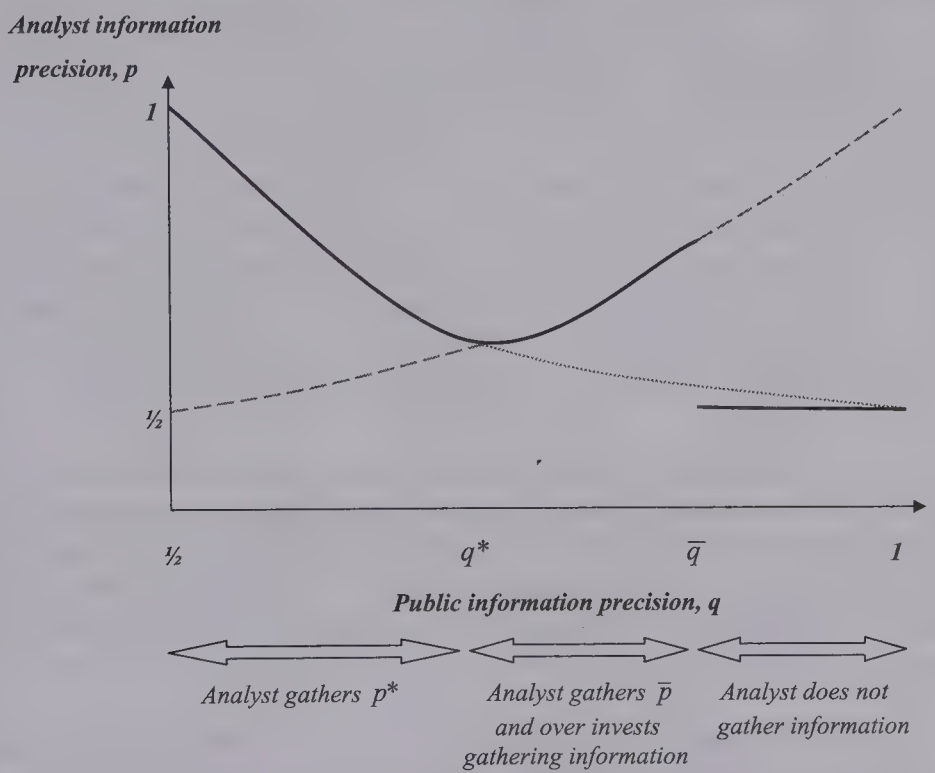
mation he gathers relative to that which he would gather if he could commit to truthfully reveal his private information. In summary, we have the following observation:

Remark 1: The precision of the information the analyst collects is decreasing in the precision of the public information, then increasing in the precision of the public information, and, at some sufficiently high level of precision of the public information, the analyst stops collecting private information; that is, there exists values q^* and $\bar{q} > q^*$ such that the equilibrium precision of the analyst's private information is decreasing in q over the range $(1/2,q^*]$, increasing in q over the range $(q^*, \bar{q}]$, and is uninformative for $q > \bar{q}$.

This characterization of the analyst's information-gathering behavior is illustrated in Figure 2. The range $(1/2,q^*]$ can be empty because $q^* = 1/2$ for some set of parameters.

Our model, which emphasizes the role of reporting credibility for understanding the information intermediary's behavior, contrasts with the voluntary disclosure models commonly examined in the accounting literature, which assume all disclosure must be truthful, but allow the informed party to decide whether to make such disclosure. These models offer reasons for why managers

FIGURE 2
Characterization of the Analyst's Information-Gathering Behavior



This figure characterizes the relation between the precision of the public information q and the analyst's equilibrium choice of precision p. The solid line reflects the analyst's equilibrium precision choice, the dashed line reflects the level of precision necessary to credibly communicate p, and the dotted line represents the level of precision chosen when credibility issues are absent, p^*.

might not voluntarily release private information. Our model suggests that analysts might not disclose non-verifiable information because, in anticipation of being unable to credibly reveal it, they abstain from gathering any information. Further, we find that analyst credibility is an important ingredient affecting the precision of information the analyst gathers. It can cause an analyst to underinvest or overinvest in information collection relative to the model in which any disclosure is always truthful.[13]

Remark 1 provides some insight into the recent evolution of the sell-side analyst industry. Recent reporting requirement changes have greatly increased the amount of information firms make available directly to investors. These changes have eroded stock analysts' information advantage. Remark 1 suggests that, in response to these changes in the information environment, some intermediaries have reduced the information they collect, others have increased the quality of their analysis in response to the heightened competition from other sources, and yet others have elected to no longer gather information on particular firms.

Remark 1 is not empirically helpful because it suggests "anything can happen." To offer more definitive guidance as to how intermediaries alter the quality of their analysis in response to changes in public information, we consider how some environmental characteristics might influence analyst behavior. In particular, we assess how the extent of the incentive conflict, captured by ϕ, and the degree of prior uncertainty about the firm's payoffs, as captured by $(s - f)$, alter analyst behavior. Doing so requires that we first examine how changes in these characteristics affect the critical values of precision.

Corollary 8: The precision of the analyst's information chosen when the analyst's information is public, p^*, is unaffected by the extent of incentive misalignment ϕ and is increasing in the difference between the state payoffs $(s - f)$; that is, $\partial p^*/\partial \phi = 0$ and $\partial p^*/\partial(s - f) > 0$. The minimum precision of the analyst's private information necessary for the analyst to credibly communicate, $\bar{p}$, is increasing in the extent of incentive misalignment ϕ and decreasing in the difference between the state payoffs $(s - f)$; that is, $\partial \bar{p} / \partial \phi > 0$ and $\partial \bar{p} / \partial (s - f) < 0$. Finally, the level of precision of the analyst's private information for which the analyst is indifferent between gathering and not gathering information, $p^{\#}$, is unaffected by the extent of incentive misalignment ϕ and is increasing in the difference between the state payoffs $(s - f)$; that is, $\partial p^{\#} / \partial \phi = 0$ and $\partial p^{\#} / \partial (s - f) > 0$.

We develop intuition for how p^* behaves in response to these parameters. The value p^* is derived assuming the analyst must report truthfully. Accordingly, the extent of the incentive misalignment between the players, ϕ, does not influence the choice of p^*. In contrast, as the prior variance of the firm's payoffs, which is proportional to $(s - f)$, increases, both players prefer more precise information, implying an increase in p^*. Because the value $p^{\#}$ presumes the analyst can credibly communicate the information gathered, $p^{\#}$ varies in a similar fashion as p^* to changes in ϕ and $(s - f)$.

We now turn to how $\bar{p}$ behaves. The analyst is more inclined to truthfully reveal his privately observed signal when the decision-maker's actions are highly responsive to his disclosure. The decision-maker is more responsive when the analyst's information is relatively more precise or the

[13] Examining the effect on the analyst behavior of changes in the precision of the other information is reminiscent of the voluntary disclosure literature that considers the credibility of a firm manager's communication when investors obtain other information that allows them to assess the veracity of the manager's disclosure. For instance, Sansing (1992) and Stocken (2000) examine managers' earnings forecasting behavior when investors can use an audited earnings report to assess a forecast's credibility.

state payoffs are more divergent. When the misalignment of incentives is large, the minimal precision of analyst information necessary for credible communication must be high. Alternatively, when the state payoffs are far apart and prior uncertainty is large, the decision-maker will be more responsive to the analyst's information. Therefore, the minimal precision of the analyst's information necessary for credible communication need not be high.

Given Corollary 8, we now specify how the extent of incentive misalignment, ϕ, influences the analyst's response to increases in the precision of the public information.

Corollary 9: If incentive misalignment is low, then the analyst will choose to gather information with precision p^*, implying the equilibrium choice of precision is decreasing in the precision of public information. Alternatively, if incentive misalignment is moderate, then the analyst will choose to gather information with precision $\bar{p}$, implying equilibrium choice of precision is increasing in the precision of public information. Finally, if incentive misalignment is high, then the analyst will abstain from collecting information.

Within the sell-side analyst context, these observations suggest that intermediaries with large incentive conflicts, such as those whose employers do a great deal of banking business (see Lin and McNichols 1998) or those who have substantial equity positions in the firm's stock, may decide to not gather private information in response to an increase in public information. In contrast, analysts that face few conflicts of interest will respond more modestly and simply reduce their analysis. Finally, analysts with more moderate conflicts of interest will be more inclined to increase their analysis in response to increases in public information.

Prior payoff uncertainty, which is proportional to $(s - f)$, also influences the response of intermediaries to the increase in public information, as formalized in Corollary 10.

Corollary 10: If the difference between the state payoffs is low, then the analyst will abstain from collecting information. Alternatively, if the difference between the state payoffs is moderate, then the analyst will choose to gather information with a level of precision $\bar{p}$, implying the equilibrium choice of precision is increasing in the precision of public information. Finally, if the difference between the state payoffs is high, then the analyst will choose to gather information with a level of precision p^*, implying the equilibrium choice of precision is decreasing in the precision of public information.

Decision-makers might be highly uncertain about a firm's payoffs, modeled as $(s - f)$ being large, when the firm is in an industry that uses an innovative and unproven technology. In response to an increase in public information, Corollary 10 predicts that firms with highly risky payoffs will face relatively modest declines in analyst attention, those with low risk payoffs will have analysts not gather any information, and those with moderately risky payoffs will receive greater analyst attention.

Decision-Maker Information

In the benchmark setting in which the analyst is constrained to report truthfully, the quality of the decision-maker's information is increasing in the quality of the public information. In the setting in which the analyst's communication is unconstrained, however, the relation between the quality of the public information and the decision-maker's information is altered by changes in the analyst's information-gathering incentives. Hence, we revisit the relation between q and the decision-maker's information.

Figure 2 shows that when the quality of public information is low, that is, $q \in (1/2, q^*]$, the analyst chooses the same level of precision p^* as in the benchmark setting with truthful reporting.

As established in the benchmark setting, p^* declines in q because the public information *substitutes* for the analyst's information. Nevertheless, the quality of the decision-maker's information is increasing in q because this increase offsets the decrease in the precision of the analyst's information.

As the quality of the public information increases above q^*, however, credibility issues arise in the unconstrained reporting setting and the analyst chooses $\bar{p}$, the lowest level of precision for which credible communication is possible. In this case, the decision-maker's quality of information is given by:

$$-E[var[\tilde{\alpha}|x,y;\bar{p}]|x] = -\bar{p}(1-\bar{p})(s-f)\Delta(\bar{p})/(2\bar{p}-1).$$

Because $\bar{p}$ is increasing in the quality of the public information, it follows that the analyst increases his precision choice in response to improvements in public information. Consequently, the public information has a *complementary* effect on the analyst's information-gathering activities.

Finally, as the precision of the public information becomes so large that the analyst's expected utility when $\bar{p}$ is chosen is less than the analyst's expected utility when no information is collected, that is, $p = 1/2$, the analyst decides not to gather any information. Since the decision-maker only observes the public information x, the decision-maker's quality of information when $p = 1/2$ is given by $-Var[\tilde{\alpha} \mid x] = -q(1-q)(s-f)^2$. This observation implies that an increase in q above some threshold value $\bar{q}$ causes the decision-maker's information quality to fall discontinuously at $\bar{q}$. Thus, the increase in public information squashes the analyst's willingness to collect private information, which makes the decision-maker worse off. Of course, as the precision of the public information continues to increase, the quality of the decision-maker's information increases and attains a maximum value at $q = 1$. This observation is formalized in Corollary 11.

Corollary 11: Suppose the analyst's precision choice is observable and the analyst need not report truthfully. As the precision of the public information increases from $q = 1/2$, the decision-maker's quality of information initially increases, falls discontinuously at $\bar{q}$ when the analyst no longer gathers private information, and then increases as q approaches 1.

Corollary 11 shows that increases in the quality of public information typically enhance the decision-maker's welfare except at a threshold level of precision $\bar{q}$, at which the public information crowds out the analyst's information. The intuition for why the quality of the decision-maker's information drops discontinuously at $\bar{q}$ is as follows. The analyst can credibly communicate only when the public information is not too precise. When the precision of the public information increases above a threshold $\bar{q}$, it crowds out the analyst's information. For public information with precision $\bar{q}$, the analyst chooses information with precision $\bar{p} > p^* > 1/2$, as illustrated in Figure 2. However, when q increases above $\bar{q}$, the decision-maker loses the analyst's information that has a level of precision $\bar{p} > 1/2$. This loss causes a *discontinuous* decrease in the quality of the decision-maker's information. While it is displaced by an increase in quality of the public information, the increase in the public information yields only a *continuous* increase in the quality of the decision-maker's information. The combination of these two effects is a discontinuous fall in the decision-maker's utility as q increases above $\bar{q}$. Further, we can show that the magnitude of the drop in the decision-maker's utility increases in the interest misalignment parameter ϕ.

Corollary 11 suggests that, at times, improving the quality of public information can crowd out the analyst's ability to credibly communicate private information. In the financial reporting environment, policymakers often decide whether information of a given precision should be gathered and disclosed rather than choosing the precision of the public information. In essence, a policymaker chooses between disclosure and no disclosure. To assess whether the crowding out

phenomenon can make *no* disclosure the policymaker's preferred choice, we analyze the difference between the decision-maker's information quality in the absence of disclosure, which is equivalent to $q = 1/2$, and the decision-maker's information quality when other information of precision q is publicly disclosed.

We show that the introduction of public information can make the decision-maker worse off within the context of an example with a cost function $c(p) = ((p - 1/2)/(1-p))^n$, where $n > 2$. Set $s = 2$, $f = 3/20$, $\phi = 1/2$, and $n = 3$. On one hand, consider the setting in which public information is absent (i.e., $q = 1/2$). The analyst chooses $\bar{p} = 1/2 + \phi/(2(s-f)) = 0.635$ because $p^* = 0.608 < \bar{p} < p^{\#} = 0.649$. Thus, the analyst gathers more precise information than he would gather if he could commit to truthfully reveal his privately observed information. Given the analyst chooses a level of precision $\bar{p}$ and communicates truthfully in the reporting game, the decision-maker's quality of information when $q = 1/2$ is given by: $-E[var[\tilde{\alpha} \mid y; \bar{p}]] = -\bar{p}(1-\bar{p})(s-f)^2 = -0.793$. On the other hand, consider an environment in which the public disclosure of information is mandatory and it has precision $q \in (0.614, 0.635)$. For instance, suppose $q = 0.625$; we then observe $p^* = 0.602 < p^{\#} = 0.640 < \bar{p} = 0.643$. In this setting, it is so costly to gather information that the analyst prefers to not gather any private information at all; that is, the analyst chooses $p = 1/2$. Although the presence of any public information crowds out the analyst's information, the decision-maker benefits from observing the public information. The decision-maker's quality of information when $q = 0.625$ is given by $-Var[\tilde{\alpha} \mid x] = -q(1-q)(s-f)^2 = -0.802$. When $q \in (0.614, 0.635)$, the benefit to the decision-maker of observing the public information is not sufficient to offset the analyst's information that it displaces. Hence, the introduction of public information reduces the total quality of the decision-maker's information and makes her worse off. Alternatively, when the precision of the public information is such that $q \in (0.5, 0.614)$, the introduction of public information encourages information gathering and leaves the decision-maker better off than when public information is absent. Last, when $q \in (0.635, 1)$, the public information is sufficiently precise that its introduction makes the decision-maker strictly better off, even though it eliminates the analyst's willingness to gather private information.

We have raised the possibility that introducing mandatory disclosures of other information can reduce the total information available to the decision-maker when the state space is binary. One might be concerned that this result can arise only when the state space is discrete, thereby allowing the analyst either to communicate truthfully or not to communicate at all. This possibility result, however, can also be established when the state space is continuous and the analyst can only partially reveal his privately observed information.[14]

The observation that additional public information may reduce the players' welfare is reminiscent of a result in Christensen's (1982) study of performance standards in an agency setting. In that study, Christensen (1982) shows that the principal may be worse-off when the players observe additional contractual information before the agent takes an action. Our setting differs from an agency setting because, in our setting, a receiver *cannot* commit (i.e., contract) to use a report in a particular way. Consequently, the character of our equilibrium and the economic forces delivering our results differ from those in Christensen (1982).

Finally, we examine the effect of the incentive misalignment ϕ and the difference between the state payoffs $(s - f)$ on the decision-maker's expected utility. Incentive misalignment affects the decision-maker's expected utility when, in equilibrium, the analyst chooses the level of precision $\bar{p}$ only. In this case, as the level of incentive misalignment increases, the precision of information the analyst must gather to credibly communicate increases, which heightens the decision-maker's expected utility. Similar to the effect of incentive misalignment, the effect of the difference be-

[14] Further details available from the authors on request.

tween the state payoffs $(s - f)$ on the decision-maker's expected utility depends on the precision of information that the analyst gathers in equilibrium. When the analyst chooses a level of precision p^*, then the effect of $(s - f)$ on $-E[var[\tilde{\alpha} \mid x, y; p^*] \mid x]$ is ambiguous. An increase in $(s - f)$ has a negative direct effect on $-E[var[\tilde{\alpha} \mid x, y; p^*] \mid x]$ and an indirect effect through p^* on $-E[var[\tilde{\alpha} \mid x, y; p^*] \mid x]$. Whether the negative direct effect dominates the indirect effect depends on the parameter values and the cost of gathering information $c(p)$. On the other hand, when the analyst gathers information with precision $\bar{p}$ or does not gather any information, then an increase in $(s - f)$ lowers the decision-maker's expected utility.

V. UNOBSERVABLE PRECISION

At times a decision-maker might be unable to determine the analyst's expertise from the stock report. Hence, we extend the unconstrained reporting setting and assume the analyst's choice of precision p is unobservable. In contrast to the analysis when p is public information, analyzing the setting in which the analyst's precision choice p is not public is more complex because the decision-maker must conjecture the analyst's choice of p and the analyst must conjecture the decision-maker's response to the analyst's report. In the equilibrium to this game, the choices of both players must maximize their objective functions conditional on each having conjectures that are consistent with the other player's equilibrium choices. As in the earlier setting in which the precision choice is observable, we begin by considering the analyst's precision choice when the analyst must report truthfully. We then address the analyst's credibility problem.

Assume the analyst must truthfully report his information y, but privately chooses its precision. Denote the decision-maker's conjecture regarding the analyst's precision choice as $\hat{p}$. When the public information x and analyst information y is realized, the decision-maker's action choice a_{xy} must satisfy:

$$a_{xg} = \frac{\hat{p}t_x}{\hat{p}t_x + (1-\hat{p})(1-t_x)}s + \frac{(1-\hat{p})(1-t_x)}{\hat{p}t_x + (1-\hat{p})(1-t_x)}f \tag{8}$$

and:

$$a_{xb} = \frac{(1-\hat{p})t_x}{(1-\hat{p})t_x + \hat{p}(1-t_x)}s + \frac{\hat{p}(1-t_x)}{(1-\hat{p})t_x + \hat{p}(1-t_x)}f, \tag{9}$$

where $t_x = q$ when $\tilde{x} = h$ and $t_x = 1 - q$ when $\tilde{x} = l$. Equations (8) and (9) imply $s \geq a_{xg} \geq a_{xb} \geq f$ in any equilibrium. Denote the analyst's conjecture regarding the decision-maker's response to realization (x,y) as $\hat{a}_{xy}$. Consider the analyst's precision choice given x and conjectured actions $s \geq \hat{a}_{xg} \geq \hat{a}_{xb} \geq f$. Given the particular realization of $\tilde{x}$, the analyst's choice of p maximizes:

$$t_x p(\phi\hat{a}_{xg} - (\hat{a}_{xg} - s)^2) + (1 - t_x)(1 - p)(\phi\hat{a}_{xg} - (\hat{a}_{xg} - f)^2) + t_x(1 - p)(\phi\hat{a}_{xb} - (\hat{a}_{xb} - s)^2)$$
$$+ (1 - t_x)p(\phi\hat{a}_{xb} - (\hat{a}_{xb} - f)^2) - c(p).$$

Because $c(p)$ is sufficiently convex, the analyst's interior choice of p satisfies the following first-order condition:

$$(\hat{a}_{xg} - \hat{a}_{xb})(t_x(2s - \hat{a}_{xg} - \hat{a}_{xb}) + (1 - t_x)(\hat{a}_{xg} + \hat{a}_{xb} - 2f) + \phi(2t_x - 1)) - c'(p) = 0. \tag{10}$$

Substituting in p for $\hat{p}$ in (8) and (9) and then taking those actions and substituting them in for $\hat{a}_{xg}$ and $\hat{a}_{xb}$ in (10) yields the following condition that any equilibrium choice for p, namely p^*_{no}, must satisfy:

$$\Delta^2(p^*_{no})/(2p^*_{no} - 1) - c'(p^*_{no}) + \Delta(p^*_{no})\phi(2t_x - 1) = 0. \tag{11}$$

Lemma 12 characterizes the analyst's precision choice.

Lemma 12: Suppose the analyst's precision choice is unobservable and the analyst must report truthfully. There always exists an equilibrium precision choice, p^*_{no}, in which the analyst does not gather any information, $p = 1/2$. In addition, if the public signal is favorable, $\tilde{x} = h$, there always exists a unique equilibrium precision choice $p^*_{no} \in (p^*, 1)$; alternatively, if the public signal is unfavorable, $\tilde{x} = l$, there may exist one or more equilibrium precision choices, all of which have the property that $p^*_{no} \in (1/2, p^*)$.

The first point worth noting in Lemma 12 is that there always exists an equilibrium in which the analyst does not provide any information. This equilibrium, which is analogous to the babbling equilibrium that always exists in a standard cheap-talk game, arises because: (1) the analyst has no incentive to gather information if the decision-maker does not respond to the resulting signal, and (2) the decision-maker has no reason to respond to a signal that is believed to be uninformative. Thus, in contrast to cheap-talk games without endogenous information collection that always feature a babbling equilibrium, in this setting in which information collection is endogenous, we find that there is always an equilibrium in which the analyst does not provide the decision-maker with any information even though the analyst's report is constrained to be truthful. The second point worth noting is that there always exists one, and perhaps more than one, equilibrium in which the analyst gathers information for at least one of the two realizations of the public signal.

To develop intuition for the equilibrium with an interior choice for p^*_{no}, we consider the analyst's reporting problem for each realization x. When $\tilde{x} = h$ and the analyst's conjecture of the decision-maker actions are $\hat{a}_{hg} > \hat{a}_{hb}$, the decision-maker can increase the probability of inducing $\hat{a}_{hg}$ by increasing precision p because it increases the likelihood that the analyst's signal y will be favorable. As a consequence, the analyst, who prefers a higher action than the decision-maker, chooses more precise information. The analyst does not choose the most precise information, however, because of the increasing marginal cost of precision. On the other hand, when $\tilde{x} = l$ and the analyst's conjecture of the decision-maker actions are $\hat{a}_{lg} > \hat{a}_{lb}$, the analyst can increase the probability of inducing the high action, $\hat{a}_{lg}$, by reducing the precision p because the likelihood that the analyst's signal will be favorable is decreasing in p. Accordingly, the analyst has an incentive to choose a lower precision p, which is mitigated by the concern that a decrease in the precision of the report increases the probability of a mismatch between the decision-maker's action and the state.

As indicated in Lemma 12, the relation between the equilibrium precision choice when the precision choice is not observable, p^*_{no}, and the precision choice when precision is observable, p^*, varies depending upon whether the public information x is favorable, $\tilde{x} = h$, or unfavorable, $\tilde{x} = l$. To develop intuition for the difference between p^* and p^*_{no}, we recognize that, when the precision choice p is not observable, the analyst chooses a level of precision without considering its impact on the decision-makers' actions. However, when p is observable, the analyst considers the impact of his precision choice on the decision-makers' actions, and the impact on those actions depends on whether $\tilde{x} = h$ or $\tilde{x} = l$. Regardless of whether $\tilde{x} = h$ or $\tilde{x} = l$, a marginal increase in precision p increases a_{xg} and decreases a_{xb}. When the public information is favorable, $\tilde{x} = h$, the magnitude of the marginal decrease in a_{hb} associated with an increase in p dominates the marginal increase in a_{hg} because the contradictory signal, $\tilde{y} = b$, has a larger influence on beliefs than the confirming signal $\tilde{y} = g$. As a consequence, the analyst is less willing to increase p when p is observable. Accordingly, when $\tilde{x} = h$, we find $p^* < p^*_{no}$. On the other hand, when $\tilde{x} = l$, the magnitude of the marginal increase in a_{lg} associated with an increase in p dominates the marginal decrease in a_{lb} because the contradictory signal, $\tilde{y} = g$, has a larger influence on beliefs. Therefore, the analyst has a greater incentive to increase p when p is observable. Accordingly, when $\tilde{x} = l$, $p^*_{no} < p^*$.

Having considered the game when the analyst must report y truthfully, we now consider how

credibility issues affect the equilibrium outcomes by assuming the analyst is not constrained to report truthfully. Suppose the favorable public signal $\tilde{x} = h$ is realized. In this case, Lemma 12 shows there exists one interior value of precision that can be an equilibrium outcome in the game in which credibility is not an issue. When credibility is also an issue, Lemma 4 implies that the value of precision p^*_{no} can be sustained as an equilibrium outcome only if $p^*_{no} \geq \bar{p}$. In contrast, if $p^*_{no} < \bar{p}$, then the analyst will always choose $p = 1/2$ and not collect any information in equilibrium. Notice that, unlike the case in which the analyst's choice is observable, there does not exist an equilibrium in which the analyst chooses $\bar{p}$. An equilibrium where the analyst choose precision $\bar{p}$ (where $\bar{p} > p^*_{no}$) cannot be sustained because the decision-maker cannot observe the analyst's actual precision choice; therefore, the analyst will choose to deviate from gathering information with precision $\bar{p}$. Alternatively, suppose the unfavorable signal $\tilde{x} = l$ is realized. If there exists an interior equilibrium in this case, then it must again be true that p^*_{no} exceeds $\bar{p}$. Note, however, that the likelihood that this condition is satisfied when $\tilde{x} = l$ is lower because p^*_{no} must be lower in the case when $\tilde{x} = l$ than in the case when $\tilde{x} = h$. Hence, the analyst is more likely to not collect any information when the public information is unfavorable. The analyst's equilibrium behavior is summarized in Proposition 13.

Proposition 13: Suppose the analyst's precision choice is unobservable and the analyst need not report truthfully. If $\bar{p} \leq p^*_{no}$, then the analyst chooses to gather information with a level of precision p^*_{no} and communicates truthfully in the reporting stage. Alternatively, if $p^*_{no} < \bar{p}$, then the analyst does not collect any private information.

The observation that an analyst's precision choice varies with firm performance provides an explanation of empirical evidence regarding analyst coverage documented by McNichols and O'Brien (1997). They find analysts are less likely to opine on poorly performing firms. The explanation they offer for this finding is that analysts avoid issuing sell recommendations so as not to jeopardize their prospects of winning investment banking business. Our model provides another explanation for the correlation they observe in the data: gathering sufficiently precise information to be able to credibly communicate is too costly for a poorly performing firm. Consequently, analysts do not gather private information about the firm.

VI. CONCLUSION

We analyze analyst information-gathering and reporting activities in a setting in which the analyst chooses the precision of information to collect in response to changes in the precision of public information about a firm. A key feature that distinguishes our model from the extant reporting literature is that the analyst is not restricted to report truthfully. Rather, recognizing credibility is an important feature of the analyst environment, we assume the analyst reports in a self-serving fashion.

We consider how the analyst's information-gathering behavior varies depending on whether the decision-maker observes the analyst's precision choice. First, when the analyst's precision choice is publicly observed, we obtain the following key results. When the analyst's reporting credibility is not an issue, an increase in the precision of public information causes the analyst to reduce the precision of information privately collected and truthfully communicated. Although the public information substitutes for the private information, an increase in the precision of the public information nevertheless makes the decision-maker better off. On the other hand, when an analyst's reporting credibility becomes salient, an increase in the precision of public information causes the analyst to collect more precise information. In this case, the public information has a complementary effect on the analyst's information gathering, which makes the decision-maker better off. If the precision of the public information continues to increase, however, then a thresh-

old is reached beyond which an analyst can no longer economically gather sufficiently precise information necessary to credibly communicate. As a consequence, the analyst discontinues gathering any private information. The analyst's decision not to gather any information makes the decision-maker worse off—which aligns with the concern that some recent regulations might make investors worse off. Indeed, we find, in some circumstances, that a decision-maker would be better off if public information was never introduced than to have public information that undermines an analyst's willingness to gather information.

Second, when investors do not observe the analyst's precision choice, we find that the precision of the analyst's information depends on the realization of the public signal. After observing favorable public information, the analyst gathers more precise information when the precision choice is not publicly observable than when it is publicly observable. Conversely, after observing unfavorable public information, the analyst chooses less precise information when the precision choice is not publicly observable relative to when it is publicly observable.

In light of recent changes in the analyst environment, our model focuses on the behavior of analysts. The model, however, is quite general. The insights might be applied to the voluntary information-gathering activities of a firm manager in the presence of changes in the mandatory reporting environment, particularly when the manager's and investors' interests are imperfectly aligned and the manager's voluntary disclosure is not verifiable—such as in the case of management earnings forecasts. The insights also might be applied to the information-gathering and communicating behavior of a seller of an experience good when the potential buyer has access to another source of unbiased information—such as *Consumer Reports*—that reveals the properties of the seller's product. Alternatively, the insights might be applied to the interaction between a venture capitalist and a firm that has private information about an innovative technology for which it is seeking venture financing.

APPENDIX

This appendix contains the proofs of the lemmas, propositions, and corollaries.

Proof of Lemma 1

After differentiating $U_i(\,.\,)$ in (3) with respect to p and using the first-order condition, the optimal choice of p^* is such that $\Delta^2(p^*)/(2p^*-1)-c'(p^*)=0$. Given U_i is a twice differentiable function, calculating $\partial^2 U_i/\partial p^2$ and evaluating the expression $\partial^2 U_i/\partial p^2$ at p^* where p^* is such that $c'(p^*)=\Delta^2(p^*)/(2p^*-1)$ yields $2c'(p^*)\Delta'(p^*)/\Delta(p^*)-c'(p^*)/(p^*-1/2)-c''(p^*)$. For a unique maximum it must necessarily be the case that $c''(p^*)/c'(p^*)>2\Delta'(p^*)/\Delta(p^*)-1/(p^*-1/2)$. Since the right-hand side of the inequality is increasing in q, the condition $c''(p^*)/c'(p^*)>(1-3p^*(1-p^*))/(p^*(1-p^*)(p^*-1/2))$ is sufficient for the objective function to be strictly quasi-concave, which the cost function $c(p)$ is assumed to satisfy. ■

Proof of Corollary 2

Applying the implicit function theorem to (5) and using the fact that $U_i(p)$ is a strictly quasi-concave and twice differentiable function that yields the observation $\partial p^*/\partial q<0$. ■

Proof of Corollary 3

To establish that $\partial(-E[var[\tilde{\alpha}\,|\,x,y;p^*]\,|\,x])/\partial q>0$, observe that:

$$\frac{\partial(-E[var[\tilde{\alpha}|x,y;p^*]|x])}{\partial q}=\frac{\partial(-E[var[\tilde{\alpha}|x,y;p^*]|x])}{\partial q}+\frac{\partial(-E[var[\tilde{\alpha}|x,y;p^*]|x])}{\partial p^*}\frac{\partial p^*}{\partial q}$$

$$\propto p^{*2}(1-p^*)^2(2q-1)+q^2(1-q)^2(2p-1)\frac{\partial p^*}{\partial q}. \qquad (12)$$

Consider $\partial p^*/\partial q$. Applying the implicit function theorem to (5) yields:

$$\frac{\partial p^*}{\partial q} = -\frac{\dfrac{\partial(\partial U_i/\partial p|_{p^*})}{\partial q}}{\dfrac{\partial(\partial U_i/\partial p|_{p^*})}{\partial p^*}} = -\frac{-\dfrac{2q(1-q)(2p^*-1)p^*(1-p^*)(2q-1)(s-f)^2}{((1-p^*)q+(1-q)p)^3(qp^*+(1-q)(1-p^*))^3}}{\dfrac{2q^2(1-q)^2(s-f)^2(3p^*(2q-1)^2(p-1)+1-3q+3q^2)}{((1-p^*)q+(1-q)p^*)^3(qp^*+(1-q)(1-p^*))^3} - c''(p^*)}. \tag{13}$$

Observe that $\partial(\partial U_i/\partial p|_{p^*})/\partial p^* < 0$ because U_i is a strictly quasi-concave and twice differentiable function. Substituting (13) into (12) and using $\partial(\partial U_i/\partial p|_{p^*})/\partial p^* < 0$ yields:

$$\frac{\partial(-E[var[\tilde{\alpha}|x,y;p^*]|x])}{\partial q} \propto p^*(1-p^*)c''(p^*) - \frac{2\Delta^2(p^*)(1-3p^*(1-p^*))}{(2p^*-1)^2}.$$

Using the assumption that $c''(p)/c'(p) > (1-3p(1-p))/(p(1-p)(p-1/2))$ and the relation $c'(p^*) = \Delta^2(p^*)/(2p^*-1)$ given in (5) yields:

$$c''(p^*) > 2\Delta^2(p^*)(1-3p^*+3(p^*)^2)/(p^*(1-p^*)(2p^*-1)^2).$$

Using this relation, which is both necessary and sufficient, yields:

$$\frac{\partial(-E[var[\tilde{\alpha}|x,y;p^*]|x])}{\partial q} \propto p^*(1-p^*)c''(p^*) - \frac{2\Delta^2(p^*)(1-3p^*(1-p^*))}{(2p^*-1)^2} > 0.$$

■

Proof of Lemma 4

The analyst's cost of gathering information is sunk and p, which is commonly observed, is fixed. If the analyst communicates truthfully, then the decision-maker's action is $a = E[\tilde{\alpha}\,|\,x, y]$. Given the decision-maker's action, the analyst reports observing g if and only if:

$$E[\phi E[\tilde{\alpha}|x,g] - (\tilde{\alpha} - E[\tilde{\alpha}|x,g])^2|g,x] \geq E[\phi E[\tilde{\alpha}|x,b] - (\tilde{\alpha} - E[\tilde{\alpha}|x,b])^2|g,x]. \tag{14}$$

This truth-telling condition is trivially satisfied because the analyst never wants to dissemble and induce a lower action than the one the decision-maker would take given the analyst's information. The analyst reveals observing b if and only if:

$$E[\phi E[\tilde{\alpha}|x,b] - (\tilde{\alpha} - E[\tilde{\alpha}|x,b])^2|b,x] \geq E[\phi E[\tilde{\alpha}|x,g] - (\tilde{\alpha} - E[\tilde{\alpha}|x,g])^2|b,x]. \tag{15}$$

This condition can be shown to be satisfied if and only if $E[\tilde{\alpha}\,|\,x, g] - E[\tilde{\alpha}\,|\,x, b] - \phi \geq 0$. Consequently, a necessary and sufficient condition for the analyst to reveal y when $\tilde{x} = l$ or for the analyst to reveal y when $\tilde{x} = h$ is:

$$E[\tilde{\alpha}|x,g] - E[\tilde{\alpha}|x,b] - \phi = \Delta(p) - \phi \geq 0. \tag{16}$$

As the incentive compatibility conditions are identical, without loss of generality, set $\tilde{x} = l$ in (16). When $q = 1/2$, (16) can be expressed as $(2p - 1)(s - f) - \phi \geq 0$. Solving for p yields $p \geq 1/2 + \phi/(2(s - f))$. On the other hand, when $q \in (1/2,1)$, solving for p in (16) yields two roots: the negative root has the feature that $p < 1/2$ and is discarded; the positive root has the feature that $p > 1/2$. Hence, the truth-telling constraint requires:

$$p \geq \bar{p} \equiv \frac{1}{2} + \frac{\sqrt{4q^2(1-q)^2(s-f)^2 + \phi^2(2q-1)^2} - 2q(1-q)(s-f)}{2\phi(2q-1)^2}. \tag{17}$$

■

The following lemma is used in the proof of Proposition 5:

Lemma 14: Let $p^* \in (1/2, \bar{p})$. The analyst will gather information with precision $p > 1/2$ only if the analyst can credibly communicate it in the reporting game.

Proof of Lemma 14

The proof proceeds by supposing to the contrary that the analyst chooses to gather information with precision $p > 1/2$ even though he cannot credibly communicate it and then establishing a contradiction. The proof is straightforward and is omitted. ■

Proof of Proposition 5

For the analyst to credibly communicate his private information y in the reporting game, it follows from (16) that p must be such that $\Delta(p) - \phi \geq 0$. Further, it follows from (5) that p^* is such that $\Delta^2(p^*)/(2p^* - 1) - c'(p^*) = 0$. It follows that $p^* \geq \bar{p}$ if and only if $\phi\Delta(p^*)/(2p^* - 1) \leq c'(p^*)$. When $p^* \geq \bar{p}$, the analyst chooses p^* and communicates credibly.

When $p^* \in (1/2, \bar{p})$, Lemma 14 establishes that the analyst will only gather information if he can credibly communicate it in the reporting game. Thus, to determine the precision of information the analyst will choose to gather, we compare the analyst's expected utility from gathering the minimum precision of information that will allow him to credibly communicate in the reporting game, i.e., $p = \bar{p}$, denoted $U_i(\bar{p})$, with the expected payoff from not gathering any private information, i.e., $p = 1/2$, denoted $U_i(1/2)$. When $p^* \in (1/2, \bar{p})$, the analyst prefers $\bar{p}$ if and only if $U_i(1/2) - U_i(\bar{p}) < 0$, or, equivalently:

$$c(\bar{p}) < \frac{q^2(2\bar{p} - 1)^2(1 - q)^2(s - f)^2}{(\bar{p}(1 - q) + (1 - \bar{p})q)((1 - \bar{p})(1 - q) + \bar{p}q)}. \tag{18}$$

It can be established that there exists a unique root, denoted $p^{\#}$, that solves the equality $c(p^{\#}) = \Delta(p^{\#})q(1 - q)(2p^{\#} - 1)(s - f)$. Hence, if $p^* \in (1/2, \bar{p})$ and if $\bar{p} \leq p^{\#}$ (or equivalently, the inequality in (18) is satisfied), then the analyst gathers more information than he would if his incentives were not misaligned with those of the decision-maker—the analyst overinvests in the quality of information gathered. Conversely, if $p^* \in (1/2, \bar{p})$ and if $p^{\#} < \bar{p}$ (or equivalently, the inequality in (18) is not satisfied), then the analyst gathers less information than he would if his incentives were not misaligned with those of the decision-maker—the analyst underinvests in the quality of information gathered.

Last, it is straightforward to show a Nash equilibrium exists in which no information is credibly communicated and the analyst sets $p = 1/2$. ■

Proof of Corollary 6

$\partial\bar{p}/\partial q > 0$ is established by partially differentiating $\bar{p}$ given in (17) with respect to q and observing that $\phi < (s - f)$ for the truth-telling constraint to be satisfied. ■

Proof of Corollary 7

Recall (7); applying the implicit function theorem to expression $U_i(p^{\#}) - U_i(1/2)$ yields:

$$\partial p^{\#}/\partial q = -\partial(U_i(p^{\#}) - U_i(1/2))/\partial q / \partial(U_i(p^{\#}) - U_i(1/2))/\partial p^{\#}.$$

It is straightforward to establish that the numerator is negative for all $q \in (1/2,1)$ and 0 otherwise. To show that the denominator is negative, we use the following arguments: First, we show that $p^{\#} > p^*$. Recall that $p^{\#}$ is such that $U_i(p^{\#}) - U_i(1/2) = 0$. It then follows from Lemma 1 that $U_i(p) > U_i(1/2)$ for all $p \in (1/2, p^*]$ and $\partial U_i(p)/\partial p < 0$ when evaluated at $p > p^*$. Consequently, it must be the case that $p^* < p^{\#}$. Second, it follows from Lemma 1 that $\partial U_i(p)/\partial p > 0$ if

and only if $p < p^*$. Because $p^* < p^{\#}$, we conclude that $\partial U_i(p) / \partial p < 0$, and hence, $\partial(U_i(p^{\#}) - U_i(1/2)) / \partial p^{\#} < 0$. Combining the above arguments yields the claim that $\partial p^{\#} / \partial q < 0$. ■

Proof of Corollary 8

The proof uses arguments offered earlier and, hence, is omitted. ■

Proof of Corollary 9

If $\phi \to 0$, then the incentive compatibility conditions (see (16)) are always satisfied. Consequently, the analyst will always truthfully reveal y for any x, that is, $\bar{p} = 1/2$. On the other hand, if $\phi \to \infty$, then the incentive compatibility conditions (see (16)) are never satisfied and the analyst cannot credibly reveal any of his private information; that is, $\bar{p} = 1$. Since $p^* \in (1/2,1)$ from Lemma 1, and $\partial \bar{p} / \partial \phi > 0$ and $\partial p^* / \partial \phi = 0$ from Corollary 8, it follows from the intermediate value theorem that there exists a threshold $\bar{\phi}$ such that for all $\phi \le \bar{\phi}$, we have $\bar{p} \le p^*$. In this circumstance, it follows from Proposition 5 that the analyst will choose precision p^*. Using Corollary 2, observe that $\partial p^* / \partial q < 0$. Again using the intermediate value theorem, there exist thresholds $\bar{\phi}$ and $\hat{\phi}$ such that for all $\phi \in (\bar{\phi}, \hat{\phi}]$, we have $p^* < \bar{p} \le p^{\#}$. In this case, it follows from Proposition 5 that the analyst will choose precision $\bar{p}$. Using Corollary 6, observe that $\partial \bar{p} / \partial q > 0$. Finally, for all $\phi > \hat{\phi}$, $p^{\#} < \bar{p}$, and it follows from Proposition 5 that the analyst chooses to not collect any private information. ■

Proof of Corollary 10

If $\sigma \equiv (s - f) \to 0$, then $\Delta(p) \to 0$ and the incentive compatibility conditions (see (16)) are never satisfied; equivalently, $\bar{p} = 1$. Since $\partial \bar{p} / \partial (s - f) < 0$ from Corollary 8, the intermediate value theorem implies there exists a threshold $\hat{\phi}$ such that for all $\sigma < \hat{\phi}$, the analyst will choose to not collect any private information; that is, he will choose $p = 1/2$. Alternatively, if $\sigma \to \infty$, then $\Delta(p) \to \infty$ and the incentive compatibility conditions (see (16)) are always satisfied; equivalently, $\bar{p} = 1/2$. Since $\partial \bar{p} / \partial (s - f) < 0$, there exists a threshold $\bar{\sigma}$ such that for all $\sigma \ge \bar{\sigma}$, we have $\bar{p} \le p^* \in (1/2, 1)$. Proposition 5 implies the analyst will choose precision p^*. Using Corollary 2, observe that $\partial p^* / \partial q < 0$. Finally, combining these arguments, it follows that for all $\phi \in [\hat{\sigma}, \bar{\sigma})$, we have $p^* < \bar{p}$. Proposition 5 implies that the analyst will choose precision $\bar{p}$. Using Corollary 6, observe that $\partial \bar{p} / \partial q > 0$. ■

Proof of Corollary 11

Lemma 3 established that $\partial(-E[var[\tilde{\alpha} \mid x, y; p^*] \mid x]) / \partial q > 0$. An analogous argument proves $\partial(-E[var[\tilde{\alpha} \mid x, y; \bar{p}] \mid x]) / \partial q > 0$. It is straightforward to show $\partial(-var[\tilde{\alpha} \mid x]) / \partial q > 0$. Hence, it remains to establish that the decision-maker's quality of information falls discontinuously at $\bar{q}$. Observe that when $c(\bar{p}) = \Delta(\bar{p}) q(1-q)(2\bar{p} - 1)(s - f)$ and $q = \bar{q}$, then $U_i(1/2) = U_i(\bar{p})$, or equivalently:

$$\phi E[\tilde{\alpha}|x] - q(1-q)(s-f)^2 = \phi E[E[\tilde{\alpha}|x,y]|x] - \frac{\bar{p}(1-\bar{p})(s-f)\Delta(\bar{p})}{(2\bar{p}-1)} - c(\bar{p}). \tag{19}$$

The law of iterated expectations, $E[\tilde{\alpha} \mid x] = E[E[\tilde{\alpha} \mid x, y] \mid x]$, implies (19) may be written as:

$$-Var[\tilde{\alpha}|x] \equiv -q(1-q)(s-f)^2 = -\frac{\bar{p}(1-\bar{p})(s-f)\Delta(\bar{p})}{(2\bar{p}-1)} - c(\bar{p}) < -\frac{\bar{p}(1-\bar{p})(s-f)\Delta(\bar{p})}{(2\bar{p}-1)}$$

$$\equiv -E[var[\tilde{\alpha}|x,y;\bar{p}]|x].$$

Hence, there exists a $q = q + \varepsilon > q^*$ for $\varepsilon > 0$ such that:

$$-Var[\tilde{\alpha}|x] < -E[var[\tilde{\alpha}|x,y;\bar{p}]|x].$$

■

Proof of Lemma 12

Because $\Delta(1/2) = 0$, it follows from Equation (11) that $p = 1/2$ is an equilibrium level of precision for $t_x = q$ or $t_x = 1 - q$. Now consider the existence of an equilibrium when $p > 1/2$. First, consider the case in which $t_x = q$; that is, $\tilde{x} = h$. Rewrite Equation (11) as:

$$(1/2)\Delta(p)^2 - c'(p)(p - 1/2) + \Delta(p)\phi(2t_x - 1)(p - 1/2) = 0.$$

The derivative of the left-hand side of Equation (11) with respect to p is:

$$\frac{2\Delta'(p)}{\Delta(p)}\left(\frac{\Delta(p)^2}{2} + \Delta(p)\phi(2t_x - 1)\left(p - \frac{1}{2}\right)\right) - \frac{1}{p - \frac{1}{2}}\left(c'(p)\left(p - \frac{1}{2}\right) - \Delta(p)\phi(2t_x - 1)\left(p - \frac{1}{2}\right)\right)$$

$$- c''(p)\left(p - \frac{1}{2}\right) - \frac{\Delta'(p)}{\Delta(p)}\Delta(p)\phi(2t_x - 1)\left(p - \frac{1}{2}\right).$$

If the equation is satisfied at p, this expression can be rewritten as:

$$2\frac{\Delta'(p)}{\Delta(p)}c'(p)\left(p - \frac{1}{2}\right) - c''(p)\left(p - \frac{1}{2}\right) - \frac{\Delta(p)^2}{2\left(p - \frac{1}{2}\right)} - \frac{\Delta'(p)}{\Delta(p)}\Delta(p)\phi(2t_x - 1)\left(p - \frac{1}{2}\right),$$

where:

$$\frac{\Delta'(p)}{\Delta(p)} = \frac{1}{2\left(p - \frac{1}{2}\right)}\frac{1 - 2p(1-p)(2t_x - 1)^2 - 2t_x(1 - t_x)}{p(1-p)(2t_x - 1)^2 + t_x(1 - t_x)}.$$

Hence, the above expression can be written as:

$$\frac{1 - 2p(1-p)(2t_x - 1)^2 - 2t_x(1 - t_x)}{p(1-p)(2t_x - 1)^2 + t_x(1 - t_x)}c'(p) - c''(p)(p - 1/2) - \frac{(1/2)\Delta(p)^2}{(p - 1/2)}$$

$$- \frac{1}{2}\frac{1 - 2p(1-p)(2t_x - 1)^2 - 2t_x(1 - t_x)}{p(1-p)(2t_x - 1)^2 + t_x(1 - t_x)}\Delta(p)\phi(2t_x - 1).$$

Because (11) holds in equilibrium, this expression can be rewritten as:

$$\left(\frac{1}{p(1-p)(2t_x - 1)^2 + t_x(1 - t_x)} - 3\right)c'(p) - c''(p)(p - 1/2)$$

$$- \left(\frac{1}{2p(1-p)(2t_x - 1)^2 + 2t_x(1 - t_x)} - 2\right)\Delta(p)\phi(2t_x - 1).$$

Noting that $p(1-p)(2t_x - 1)^2 + t_x(1 - t_x)$ is decreasing in t_x and that $2t_x - 1 > 0$ because $t_x = q > 1/2$ implies that the above expression is strictly less than $[(1 - 3p(1-p)) / (p(1-p))]c'(p) - c''(p)(p - 1/2)$, which is proportional to $(1 - 3p(1-p))c'(p) - p(1-p)(p - 1/2)c''(p)$. Hence, there exists a unique equilibrium in which $p \in (1/2,1)$ if $(1 - 3p(1-p)) / (p(1-p)(p - 1/2)) \leq c''(p) / c'(p)$ for all p, which is true by assumption.

Second, consider the case of $t_x = 1 - q$; that is, $\tilde{x} = l$. There may or may not exist an equilib-

rium in which $p \in (1/2,1)$ and, sometimes there may exist more than one equilibrium in which $p \in (1/2,1)$. The proof follows from an example in which there is no equilibrium when $p \in (1/2,1)$, one equilibrium when $p \in (1/2,1)$, and two equilibria when $p \in (1/2,1)$. For the examples, let $c(p) = ((p - 1/2)/(1-p))^3$ and $q = .85$. If $\phi = 2$ and $s - f = 2$, then there does not exist a $p \in (1/2,1)$ that satisfies the equilibrium condition. If $\phi = 1$ and $s - f = 2$, then $p = .5227$ is the only $p \in (1/2,1)$ that satisfies the equilibrium condition. If $\phi = 110$, $s - f = 150$, then $p = .9072$ and $p = .5584$ both satisfy the equilibrium condition. ■

Proof of Proposition 13

Consider the case in which $\bar{p} \leq p^*_{no}$. When the analyst's information y is privately observed, it follows from Lemma 4 that the analyst can only credibly reveal it when it has a level of precision p such that $p \geq \bar{p}$. Therefore, if $\bar{p} \leq p^*_{no}$, then the analyst can credibly reveal his privately observed information with precision p^*_{no}. Alternatively, consider the case in which $p^*_{no} < \bar{p}$. It follows from Lemma 4 that the analyst cannot credibly reveal information with precision p^*_{no}. Thus, the analyst either will choose $p = 1/2$ and not gather any privately observed information, or choose $p = \bar{p}$ and credibly reveal his privately observed information. When the analyst's precision choice is unobservable, Proposition 12 implies that p^*_{no} is the only unobservable choice of precision that can be supported in equilibrium. Thus, the analyst would prefer to deviate from any unobservable precision choice $p \neq p^*_{no}$. Since $p^*_{no} < \bar{p}$ and the decision-maker cannot observe the precision choice $\bar{p}$, it follows that the analyst will not choose to gather information with precision $\bar{p}$. Since gathering information y is costly, Lemma 14 implies the analyst will choose $p = 1/2$. ■

REFERENCES

Alonso, R., W. Dessein, and N. Matouschek. 2008. When does coordination require centralization? *The American Economic Review* 98 (1): 145–179.

Arya, A., and B. G. Mittendorf. 2007. The interaction among disclosure, competition between firms, and analyst following. *Journal of Accounting and Economics* 43 (2–3): 321–339.

Aumann, R. J., and S. Hart. 2003. Long cheap talk. *Econometrica* 71 (6): 1619–1660.

Austen-Smith, D. 1993. Interested experts and policy advice: Multiple referrals under open rule. *Games and Economic Behavior* 5 (1): 3–43.

———. 1994. Strategic transmission of costly information. *Econometrica* 62 (4): 955–963.

Battaglini, M. 2004. Policy advice with imperfectly informed experts. *Advances in Theoretical Economics* 4 (1).

Bushee, B. J., D. Matsumoto, and G. Miller. 2003. Open versus closed conference calls: The determinants and effects of broadening access to disclosure. *Journal of Accounting and Economics* 34 (1–3): 149–180.

Christensen, J. 1982. The determination of performance standards and participation. *Journal of Accounting Research* 20 (2): 589–603.

Crawford, V. P., and J. Sobel. 1982. Strategic information transmission. *Econometrica* 50 (6): 1431–1451.

Dugar, A., and S. Nathan. 1995. The effect of investment banking relationships on financial analyst's earnings forecasts and investment recommendations. *Contemporary Accounting Research* 12 (1): 131–160.

Dye, R. A. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23 (1): 123–145.

Einhorn, E. 2007. Voluntary disclosure under uncertainty about the reporting objective. *Journal of Accounting and Economics* 43 (2–3): 245–274.

———, and A. Ziv. 2008. Intertemporal dynamics of corporate voluntary disclosures. *Journal of Accounting Research* 46 (3): 567–589.

Farrell, J., and R. Gibbons. 1989. Cheap talk with two audiences. *The American Economic Review* 79 (5): 1214–1223.

Gigler, F. 1994. Self-enforcing voluntary disclosures. *Journal of Accounting Research* 32 (2): 224–240.

Grossman, G. M., and E. Helpman. 2001. *Special Interest Politics*. Cambridge, MA: The MIT Press.

Hayes, R. M. 1998. The impact of trading commission incentives on analysts' stock coverage decisions and earnings forecasts. *Journal of Accounting Research* 36 (2): 299–320.

Heflin, F., K. R. Subramanyam, and Y. Zhang. 2003. Regulation FD and the financial information environment: Early evidence. *The Accounting Review* 78 (1): 1–37.

Hughes, J. S., and S. Pae. 2004. Voluntary disclosure of precision information. *Journal of Accounting and Economics* 37 (2): 261–289.

Krishna, V., and J. Morgan. 2001. A model of expertise. *The Quarterly Journal of Economics* 116 (2): 747–775.

———, and ———. 2004. The art of conversation: eliciting information from experts through multi-stage communication. *Journal of Economic Theory* 117 (2): 147–179.

Lin, H.-W., and M. F. McNichols. 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25 (1): 101–127.

Mayew, W. J. 2008. Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research* 46 (3): 627–659.

McNichols, M., and P. C. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* 35 (Supplement): 167–199.

Michaely, R., and K. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12 (4): 653–686.

Morgan, J., and P. C. Stocken. 2003. An analysis of stock recommendations. *The RAND Journal of Economics* 34 (1): 183–203.

———, and ———. 2008. Information aggregation in polls. *The American Economic Review* 98 (3): 864–896.

Newman, P., and R. Sansing. 1993. Disclosure policies with multiple users. *Journal of Accounting Research* 31 (1): 92–112.

Pae, S. 1999. Acquisition and discretionary disclosure of private information and its implications for firms' productive activities. *Journal of Accounting Research* 37 (2): 465–474.

Penno, M. 1996. Unobservable precision choices in financial reporting. *Journal of Accounting Research* 34 (1): 141–149.

Razaee, Z. 2007. *Corporate Governance Post-Sarbanes-Oxley: Regulations, Requirements, and Integrated Processes*. Hoboken, NJ: John Wiley & Sons.

Sansing, R. C. 1992. Accounting and the credibility of management forecasts. *Contemporary Accounting Research* 9 (1): 33–45.

Schack, J. 2003. Unsettled on Wall Street. *The Institutional Investor* (October): 1.

Stocken, P. C. 2000. Credibility of voluntary disclosure. *The RAND Journal of Economics* 31 (2): 359–374.

Verrecchia, R. E. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5 (1): 179–194.

———. 1990. Information quality and discretionary disclosure. *Journal of Accounting and Economics* 12 (4): 365–380.

Wolinsky, A. 2002. Eliciting information from multiple experts. *Games and Economic Behavior* 41 (1): 141–160.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 2011–2046

American Accounting Association
DOI: 10.2308/accr.2010.85.6.2011

An Empirical Analysis of Auditor Independence in the Banking Industry

Kiridaran Kanagaretnam
McMaster University

Gopal V. Krishnan
Lehigh University

Gerald J. Lobo
University of Houston

ABSTRACT: We examine auditor independence in the banking industry by analyzing the relation between fees paid to auditors and the extent of earnings management through loan loss provisions (LLP). We also examine whether this relation differs across large banks whose managements are required under the Federal Deposit Insurance Corporation Improvement Act to evaluate internal control over financial reporting and whose auditors must attest to the effectiveness of such internal controls, and small banks that are not subject to those requirements. We find that unexpected auditor fees are unrelated to earnings management for large banks. For small banks, we find greater earnings management via under-provisioning of LLP by banks that pay higher unexpected total and nonaudit fees to the auditor. These results suggest that auditor fee dependence on the audit client is associated with earnings management via abnormal LLP and is a potential threat to auditor independence for small banks. Our findings are relevant to policymakers contemplating new regulations in light of the recent banking crisis.

Keywords: *auditor independence; earnings management; auditor fees; bank loan loss provision; FDICIA.*

Data Availability: *All data presented in this paper are available from public sources.*

We thank James Bierstaker, Jeff Chen, Chris Jones, Steven Kachelmeier (editor), Sok-Hyon Kang, Krishna Kumar, Ying Li, Lihong Liang, Jim Largay III, Robert Mathieu, Erin Moore, Nandu Nayar, Dan Neeley, Mike Peters, H. Sami, Bin Srinidhi, Mary Sullivan, Scott Whisenant, Bill Zhang, seminar participants at Brock University, The George Washington University, Hong Kong Polytechnic University, Lehigh University, McMaster University, Nanyang Technological University, University of Queensland, the 2007 American Accounting Association Annual Meeting, and the 2007 Annual Meeting of the Canadian Academic Accounting Association, and two anonymous reviewers for their helpful suggestions. Professors Kanagaretnam and Lobo thank the Social Sciences and Humanities Research Council of Canada (SSHRC) for its financial support.

Editor's note: Accepted by Steven Kachelmeier, with thanks to John Core for serving as editor on a previous version.

Submitted: January 2007
Accepted: April 2010
Published Online: November 2010

I. INTRODUCTION

Auditor independence is vital to maintaining public confidence in capital markets and to the integrity of corporate financial statements. The objective of this study is to examine auditor independence in the banking industry. Banks represent a large percentage of the total public equity market and are critical to the functioning of the economy as a whole. Nevertheless, as noted by Fields et al. (2004), accounting researchers have done little to investigate the various relationships that exist between banks and their auditors despite the economic importance of the banking industry. Specifically, we provide empirical evidence on the relation between fees paid to auditors of banks and the extent of earnings management via loan loss provisions (LLP). Our study is timely and relevant given the recent banking crisis and that governments around the world are contemplating new banking regulations. In particular, our research informs policymakers on the relationship that existed between fees paid to auditors and the extent of earnings management in banks prior to the current banking crisis.

The banking industry offers a unique context to study auditor independence for a number of reasons. First, banks are subject to the scrutiny of the Federal Deposit Insurance Corporation (FDIC), the Federal Reserve Board, and other government agencies. Whether this intense regulatory oversight enhances auditor independence is clearly an important and relevant question. Second, the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA) imposed new auditing, corporate reporting, and governance reforms on depository institutions with assets exceeding $500 million (increased to $1 billion in 2005) and on their auditors (Murphy 2004). Section 36 of the Federal Deposit Insurance Act (FDI Act) and Part 363 of the FDIC's regulations require the management of these institutions to evaluate internal control over financial reporting and the auditor to attest to the effectiveness of internal controls over financial reporting.[1] Given these requirements, whether auditor independence is greater for larger banks that are subject to greater regulatory oversight relative to smaller banks is an interesting and important question.[2]

Third, bank LLP is well-suited to studying earnings management for the following reasons. Bank managers have considerable discretion in estimating LLP. This discretion allows them flexibility in using LLP for income-increasing or income-decreasing earnings management, or for smoothing earnings.[3] LLP is also by far the largest and most important accrual for banks, thus affording bank managers wide latitude in its use.[4] Prior research indicates that banks use LLP to manage earnings (Wahlen 1994; Kanagaretnam et al. 2003; Kanagaretnam et al. 2004). To the extent that banks can leverage "fee dependence" to influence their auditors to accept abnormal LLP, examining the relation between abnormal LLP and auditor fees is likely to reveal such a linkage. We reason that abnormal LLP is a better proxy for earnings management than the abnor-

---

[1] There is some prior research on FDICIA. For example, Altamuro and Beatty (2010) examine several earnings quality measures prior to and following FDICIA and find that the mandated internal control requirements increased the validity of LLP, earnings persistence, and cash flow predictability, and reduced benchmark-beating and accounting conservatism for affected versus unaffected banks.

[2] In addition to bank asset size, we also take into consideration the banks that have to comply with Section 404 of the Sarbanes-Oxley Act (SOX), which became effective in 2004, when classifying banks as large and small. We consider small banks as banks that are exempted from both FDICIA and Section 404 of SOX. We use the terms small banks and exempted banks interchangeably in our paper. Section II provides a more detailed description of the institutional details.

[3] Prior research suggests several motives for bank managers' discretionary behavior with respect to LLP including, signaling, capital management, management compensation, and income smoothing (Wahlen 1994; Collins et al. 1995; Kanagaretnam et al. 2003; Kanagaretnam et al. 2004). Income-increasing abnormal LLP may be related to incentives associated with meeting or just beating earnings benchmarks, management compensation, and income smoothing. Income-decreasing abnormal LLP may be related to incentives associated with signaling, capital management, and income smoothing.

[4] The mean (median) ratio of LLP to earnings before LLP is 17 percent (14.6 percent) for our sample banks.

mal accrual measures used in prior research.[5] In particular, our study mitigates error in measuring managerial discretion by focusing on a single accrual and a single industry. Focusing on a single accrual facilitates a sharper separation into its normal (nondiscretionary) and abnormal (discretionary) components. We use a number of industry-specific variables to better isolate the normal LLP from the abnormal LLP. Also, focusing on a single, relatively homogeneous industry provides control over other determinants of cross-sectional differences in accruals, thus increasing the reliability of the inferences from our empirical analysis.

Our main prediction is that, relative to large banks, small banks that pay higher fees to auditors are likely to engage in higher levels of income-increasing earnings management—i.e., higher fees are likely to facilitate income-increasing LLP (under-provisioning of LLP). Income-increasing LLP is of particular importance in the context of auditor independence because auditors are unlikely to be sued for income-decreasing accruals (St. Pierre and Anderson 1984). Additionally, we expect the magnitude of income-decreasing LLP to decrease for small banks as banks pay higher fees.

A possible alternative hypothesis to the impairment of auditor independence through fees is that audit fees may also reflect audit risk that is captured by the abnormal LLP. Auditors are likely to charge higher audit fees to firms that are more difficult to audit and, if firms that are more difficult to audit have higher incentives to manage earnings, this will be manifested in a positive relationship between audit fees and abnormal LLP. We attempt to address this concern by including several variables that reflect differences in audit risk across banks as well as examining the relationship using abnormal fees after explicitly taking account of possible factors driving the normal fees.

Our sample consists of 1,740 bank-year observations over the years 2000–2006. Following prior research, we estimate abnormal LLP as residuals from a regression of LLP on beginning loan loss allowance, total loans outstanding, changes in total loans outstanding, net loan charge-offs, beginning balance of non-performing loans, change in non-performing loans, loan mix, and controls for years (Wahlen 1994; Kanagaretnam et al. 2004). Next, we estimate a regression of abnormal LLP on fee measures, an indicator variable for small banks (with total assets of less than $500 million or less than $1 billion effective 2005 or non-accelerated filers under Section 404 of Sarbanes-Oxley Act [SOX] effective 2004) that are exempt from internal control regulation, the interaction of the fee measure and the indicator variable for small banks, auditor type (Big 5 versus non-Big 5 auditor), and several control variables. We use various measures of fees to capture the auditor's economic dependence on the client, including total fees, nonaudit fees, audit fees, ratio of nonaudit fees to total fees (fee ratio), abnormal (unexpected) total fees, abnormal nonaudit fees, abnormal audit fees, and abnormal fee ratio. However, our main focus is on unexpected total fees and unexpected nonaudit fees because these fee measures are based on controlling for bank-level determinants of normal fees and prior research suggests that they are better measures for examining fee dependence (Kinney and Libby 2002; Srinidhi and Gul 2007).

We report several key findings. First, for large banks, we document that both unexpected total fees and unexpected nonaudit fees have no association with either income-increasing or income-decreasing abnormal LLP. Second, in contrast to the results for large banks, we find a strong negative association between income-increasing (negative) abnormal LLP and both unexpected total fees and unexpected nonaudit fees for small banks. In other words, higher auditor fees are associated with higher income-increasing earnings management for small banks. We also docu-

[5] Prior research argues that proxies of abnormal accruals commonly used to detect earnings management are subject to serious measurement error (Guay et al. 1996; McNichols 2000, 2002; Jones et al. 2008). For example, McNichols (2002) questions the construct validity of a proxy based on aggregate accruals because of the complexity associated with modeling the estimation errors in aggregate accruals.

ment a negative association between unexpected nonaudit fees and income-decreasing (positive) abnormal LLP for small banks. These results are consistent with greater earnings management by small banks that pay higher abnormal fees to the auditor. Our results are robust to a variety of sensitivity tests, including separate subsample analyses for small and large banks, deletion of influential observations, and use of alternative models for estimating abnormal LLP.

Our findings are consistent with economic bonding between the auditor and the bank being associated with income-increasing earnings management through LLP for small banks. In particular, our results suggest that auditor fee dependence on the audit client can potentially threaten auditor independence, particularly among banks that are not subject to the same level of regulatory scrutiny as large banks.[6]

The next section describes the institutional background on FDICIA and SOX, particularly how these regulations affect small and large banks. Section III describes our research design, including the empirical models used to estimate abnormal LLP, unexpected fees paid to auditors, and tests of the relation between these variables. We also discuss how the differences in regulation of small and large banks could impact auditor independence. Section IV details the sample selection, Section V discusses the results, and Section VI provides the conclusions of the study.

II. INSTITUTIONAL BACKGROUND

In response to the breakdown of thrifts and banks during the late 1980s and early 1990s, the U.S. Congress enacted FDICIA in 1991 to strengthen the financial condition of the banking and thrift industries. FDICIA, which created deposit insurance and the FDIC, is the most important banking legislation in the U.S. since the Banking (Glass-Steagall) Act of 1933 (Benston and Kaufman 1997). Further, a decade later, Congress passed another landmark legislation, the Sarbanes-Oxley Act of 2002 (SOX) which, although modeled after FDICIA, is much broader in scope.[7] While FDICIA contains much more than deposit insurance reform, of particular interest are the requirements for annual audit and reporting of management's assessment of the effectiveness of internal control that are intended to aid early detection of problems in the financial management of insured banks. In June 1993, the FDIC enacted Part 363 titled, "Annual Independent Audits and Reporting Requirements," which requires each insured depository institution with $500 million or more in total assets at the beginning of the year to prepare (see Sections 363.2 and 363.3 of FDIC 1993):[8,9]

1. Audited annual financial statements;
2. A management report signed by the CEO and the CFO that contains:
   (a) a statement of management's responsibilities for:
      i. preparing financial statements,
      ii. establishing and maintaining an adequate internal control structure over financial reporting, and
      iii. complying with the laws and regulations that are designated by the FDIC and other federal banking agencies;

---

[6] Interestingly, over 85 percent of these small banks are audited by non-Big 5 auditors, suggesting that fee dependence may be greater for the non-Big 5 auditors relative to the Big 5 auditors.

[7] See Altamuro and Beatty (2010) for a mapping of sections of FDICIA with SOX.

[8] Section 112 of FDICIA added Section 36, "Early Identification of Needed Improvements in Financial Management," to the FDI Act. Section 36 gives the FDIC discretion to select an asset-size threshold of $150 million or more to exempt smaller depository institutions from the above reporting requirements. This exemption was intended to ease compliance cost for small and community banks.

[9] In addition, Section 363.5 requires each depository institution to establish an independent audit committee, the members of which shall be outside directors who are independent of the management.

(b) assessments by management of:
   i. the effectiveness of the internal control structure over financial reporting, and
   ii. the institution's compliance with the designated laws and regulations during the year;
3. The independent public accountant's report on the audited financial statements; and
4. The independent public accountant's attestation report on the assertion of management concerning the institution's internal control structure and procedures for financial reporting.

Effective 2005, the threshold for requirements 2(b)i and 4 was increased from $500 million to $1 billion in total assets (FDIC 2005b).[10] Although depository institutions with $500 million or more but less than $1 billion in total assets are currently exempt from internal control assessments by management and external auditors per Section 36, Section 404 of SOX mandates that all public companies or subsidiaries of public companies must file a report on the auditor's attestation of the effectiveness of internal control over financial reporting. Currently, Section 404 exempts non-accelerated filers—i.e., firms with public float less than $75 million.

In summary, publicly held banks with less than $500 million in total assets during the years 2000 through 2004 are exempt from Section 36 of the FDI Act. Effective 2004, publicly held banks that are non-accelerated filers are exempt from Section 404 of SOX. Currently, publicly held banks with less than $1 billion in total assets and with a public float less than $75 million are exempt from both Section 36 as well as Section 404. Privately held banks are exempt from Section 404; further, effective 2005, privately held banks with $500 million or more but less than $1 billion in total assets (or less than $500 million in total assets during the years 2000 through 2004) are exempt from Section 36 pertaining to management's and auditor's assessments of the effectiveness of internal control over financial reporting. We do not focus on privately held banks, as information on fees paid to the auditor is not disclosed for these banks.

In addition to FDICIA and SOX, banks are expected to comply with guidance issued by the SEC and the Federal Financial Institutions Examination Council (FFIEC) on LLP.[11] Both SAB No. 102 and the policy statement emphasize the methodology used to measure LLP and the documentation, rather than external financial reporting (Phillips and Lierley 2001/2002).[12]

III. RESEARCH DESIGN

In this section, we first discuss our conceptual model, depicted in Figure 1, on the relation between economic bonding and earnings management. Second, we describe the empirical model used to estimate abnormal LLP. Third, we describe the empirical model used to estimate unexpected (abnormal) measures of fees paid to the auditor. Fourth, we discuss the empirical model used to test possible linkages between abnormal LLP and unexpected fees paid to auditors.

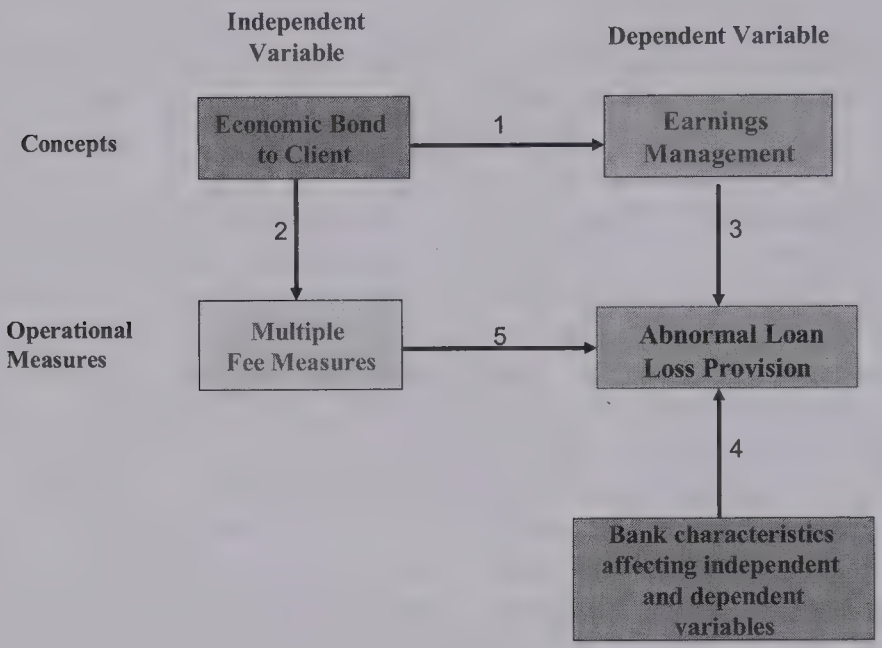
The model in Figure 1 is based on the predictive validity model of Kinney and Libby (2002) that was developed by Runkel and McGrath (1972) and Libby (1981). As in prior research, the economic bond to the audit client and earnings management by the client (link 1) are the two theoretical constructs we examine. We identify proxies to measure the economic bond and earn-

[10] Note that the FDIC increased the asset size threshold for internal control assessments by management and external auditors but retained the financial statement audit and other reporting requirements for all institutions with $500 million or more in total assets to preserve safety and soundness in financial management (FDIC 2005a).

[11] In July 2001, the SEC issued SAB No. 102, *Selected Loan Loss Allowance Methodology and Documentation Issues*, and the FFIEC issued, *Policy Statement on Allowance for Loan and Lease Losses Methodologies and Documentation for Banks and Savings Institutions*.

[12] Although this guidance applies equally to all institutions, the policy statement suggests that small banks may produce fewer supporting documents in support of their loan loss methodologies than large banks (FDIC 2001).

FIGURE 1
Conceptual Model



ings management because these two constructs are unobservable. We use multiple fee measures to capture the economic bond (link 2). Kinney and Libby (2002) reason that the economic bond is likely to strengthen as *unexpected* nonaudit and audit fees increase because the abnormal rents earned by the auditor increase. We use unexpected total fees and unexpected nonaudit fees as primary measures of the economic bond (link 2) and the abnormal component of a bank's LLP as the measure of earnings management (link 3). We examine banking firms because they provide a setting with a stronger link 3 than that characterizing prior research. By using an industry-specific measure of earnings management, we can also improve upon prior research in terms of a sharper separation between abnormal and normal accruals. We control for several bank characteristics that can also affect LLP (link 4). We then examine the relation between the various fee measures and abnormal LLP (link 5) to draw inferences about link 1.

We expect banks that are exempt from assessments of the effectiveness of internal control over financial reporting by managers and auditors to pose a greater risk to auditor independence than banks that are not exempt. When bank managers and the auditor know that the bank will not be subject to regulatory scrutiny, such a setting likely encourages earnings management by managers and could increase auditors' tolerance for earnings management relative to banks that are expected to comply with the requirements of FDICIA and Section 404 of SOX.[13]

Estimation of Abnormal Loan Loss Provisions

In this section we describe our two-stage approach to examine link 5 (see Figure 1). We estimate the normal or nondiscretionary component of LLP by regressing LLP on beginning loan

[13] Prior research suggests that increased legal liability increases the auditor's effort and audit accuracy (Boritz and Zhang 1997; De and Sen 2002). Similarly, Kim et al. (2008) argue that legal environment plays an important role in determining auditor effort.

loss allowance, beginning non-performing loans, change in non-performing loans, net loan charge-offs, change in total loans outstanding, total loans outstanding, loan mix, and controls for period effects using the following model (all variables are scaled by beginning total assets):[14,15]

$$LLP = \alpha_0 + \alpha_1 BEGLLA + \alpha_2 BEGNPL + \alpha_3 CHNPL + \alpha_4 LCO + \alpha_5 CHLOANS + \alpha_6 LOANS + LOANCATEGORIES + YEARCONTROLS + \epsilon \quad (1)$$

where:

LLP = provision for loan losses;
BEGLLA = beginning loan loss allowance;
BEGNPL = beginning non-performing loans;
CHNPL = change in non-performing loans;
LCO = net loan charge-offs;
CHLOANS = change in total loans outstanding;
LOANS = total loans outstanding; and
LOANCATEGORIES = amount of commercial loans (*COMM*), consumer loans (*CON*), real estate loans (*RESTATE*), agriculture loans (*AGRI*), loans to foreign banks and governments (*FBG*), and loans to other depository institutions (*DEPINS*).

The residuals from model (1) are the abnormal component of LLP, referred to as *ALLP*. We expect a negative coefficient on *BEGLLA* (i.e., accumulated LLP less write-offs at the beginning of the year), as a higher initial loan loss allowance will require a lower LLP in the current period. Consistent with prior research, we expect α_2, α_3, α_4, and α_6 to be positive. Higher levels of non-performing loans indicate that problems with the loan portfolio will require higher loss provisions. Therefore, we expect the beginning balance of non-performing loans (*BEGNPL*) to be positively related to LLP. We expect change in non-performing loans (*CHNPL*) to be positively related to LLP because an increase in non-performing loans will require a higher loan loss provision in the current period. We expect net loan charge-offs (*LCO*) to be positively related to LLP. As explained by Beaver and Engel (1996), because current loan charge-offs can provide information about future loan charge-offs, they can influence expectations of the collectability of current loans and hence current LLP. We expect the level of loans (*LOANS*) to be positively related to LLP because a higher level of loans will also require higher provisions. We do not offer a prediction for α_5 because the effect of change in total loans (*CHLOANS*) on LLP depends on the quality of incremental loans.[16]

Although non-performing loans and loan charge-offs serve as measures of risk, we include six additional variables to control for differences in loan composition that also likely contribute to differences in risk. For example, banks with a higher proportion of commercial and real estate loans are likely to have higher loan loss provisions than banks with a higher proportion of consumer loans. Failure to account for these differences makes the residuals from the loan loss provision model a function of the bank type. This, in turn, could affect the inferences of our auditor independence tests because both audit fees and *ALLP* may systematically vary across banks based on their loan portfolio mix.

[14] These variables have also been used in several prior studies (e.g., Wahlen 1994; Kanagaretnam et al. 2004) to estimate the normal component of LLP.
[15] We also use beginning total loans as an alternate deflator. Our results are not sensitive to this choice of deflator.
[16] For convenience, we summarize the variable definitions in Table 1.

TABLE 1

Definitions of Variables

LAFEE = natural log of audit fees;
LTOTFEE = natural log of total fees;
LNAFEE = natural log of nonaudit fees;
FEERATIO = nonaudit fees divided by total fees;
UTOTFEE = unexpected total fees from model (2);
UNAFEE = unexpected nonaudit fees from model (2);
UAFEE = unexpected audit fees from model (2);
UFEERATIO = unexpected fee ratio from model (2);
LLP = provision for loan losses divided by beginning total assets;
ALLP = abnormal LLP (residual from model (1));
LOANS = total loans outstanding divided by beginning total assets;
BEGLLA = beginning loan loss allowance divided by beginning total assets;
CHLOANS = change in total loans outstanding divided by beginning total assets;
BEGNPL = beginning non-performing loans divided by beginning total assets;
CHNPL = change in non-performing loans divided by beginning total assets;
LCO = net loan charge-offs divided by beginning total assets;
COMM = commercial loans divided by beginning total assets;
CON = consumer loans divided by beginning total assets;
RESTATE = real estate loans divided by beginning total assets;
AGRI = agriculture loans divided by beginning total assets;
FBG = loans to foreign banks and governments divided by beginning total assets;
DEPINS = loans to other depository institutions divided by beginning total assets;
BIG5 = indicator variable that equals 1 if audited by a Big 5 firm, and 0 otherwise;
MB = market-to-book ratio at the end of the year;
LMVE = natural log of market value of common equity at the end of the year;
LOSS = indicator variable that equals 1 if net income <0, and 0 otherwise;
PASTLLP = prior year's LLP divided by beginning total assets;
EBP = net income before extraordinary items and LLP divided by beginning total assets;
TIER1 = tier 1 risk-adjusted capital ratio at the beginning of the year;
TCAP = total risk-adjusted capital ratio at the beginning of the year;
LASSETS = natural log of total assets;
SECURITIES = [1 − (total securities/total assets)];
SNPL = non-performing loans divided by lagged total loans;
INTANG = intangible assets divided by total assets;
EFFICIENCY = total operating expenses divided by total revenues (measured as net interest income plus non-interest income);
SLCO = net loan charge-offs divided by loan loss allowance;
SCOMM = total commercial and agricultural loans divided by total loans;
SCON = total consumer loans divided by total loans;
SRESTATE = total real estate loans divided by total loans;
EXEMPT = banks exempted from internal control regulations. An indicator variable that equals 1 if a bank has less than $500 million in total assets during the years 2000 through 2003 or less than $500 million in total assets and public float less than $75 million for the year 2004 or less than $1 billion in total assets and public float less than $75 million for the years 2005 and 2006, and 0 otherwise;
SMALL = indicator variable that equals 1 if a bank has less than $500 million in total assets during the years 2000 through 2003; and
NF = indicator variable that equals 1 if a bank has less than $500 million in total assets and public float less than $75 million for the year 2004 or less than $1 billion in total assets and public float less than $75 million for the years 2005 and 2006, and 0 otherwise.

The six loan portfolio composition variables included in the model are commercial loans (*COMM*), consumer loans (*CON*), real estate loans (*RESTATE*), agriculture loans (*AGRI*), loans to foreign banks and governments (*FBG*), and loans to other depository institutions (*DEPINS*). We also include six year-indicator variables representing years 2000 through 2005 in model (1) to control for period-specific effects.

Next, we describe the model for computing unexpected (abnormal) fees. Auditor fees are also influenced by various economic determinants including size and complexity of audit task. Therefore, controlling for the economic determinants of auditor fees and using unexpected (abnormal) auditor fees is likely a better measure of auditor fee dependence than using actual fees.

Estimation of Unexpected Fees

Kinney and Libby (2002) argue that unexpected fees are a better measure of the auditor-client economic bond because they reflect the "excess" profit derived from an audit client. We estimate unexpected fees as the difference between actual and expected (normal) fees, where expected fees are the predicted values from a regression of auditor fees on a set of firm characteristics.

Prior research models auditor fees as a function of a firm's auditor choice, audit complexity, and audit risk, in addition to other variables (Firth 1997; Ashbaugh et al. 2003). Fields et al. (2004) examine the determinants of audit fees in the banking industry. We use the variables identified in the Fields et al. (2004) study to estimate unexpected fees. Audit fees and other fees are likely to be higher when the auditor is a Big 5 auditor. Auditor size also proxies for client size. We measure firm size as the natural log of total assets. The normal audit fee is directly related to a bank's credit risk, operating risk, liquidity risk, and capital risk. We include *SNPL, SLCO*, *SCOMM*, *SCON*, and *SRESTATE* as proxies for a bank's credit risk and the efficiency ratio (*EFFICIENCY*) as a proxy for operating risk. We measure the efficiency ratio as the ratio of total operating expenses to total revenues. As in Fields et al. (2004), we use *SECURITIES* as a proxy for liquidity risk, and intangible assets (*INTANG*) and total capital ratio (*TCAP*) to account for capital risk. In addition, we include a control for exempted banks (*EXEMPT*), which are expected to pay lower fees to auditors. We estimate the following model:[17]

$$\begin{aligned} FEE = {} & \delta_0 + \delta_1 BIG5 + \delta_2 LASSETS + \delta_3 SECURITIES + \delta_4 SNPL + \delta_5 LOSS + \delta_6 INTANG \\ & + \delta_7 EFFICIENCY + \delta_8 SLCO + \delta_9 SCOMM + \delta_{10} SCON + \delta_{11} SCESTATE + \delta_{12} TCAP \\ & + \delta_{13} EXEMPT + YEARCONTROLS + \varepsilon \end{aligned} \tag{2}$$

where:

FEE = fees paid to the auditor (*LTOTFEE* = ln (audit fees + nonaudit fees); or *LNAFEE* = ln (nonaudit fees); or *LAFEE* = ln (audit fees)); or *FEERATIO* = nonaudit fees/total fees);
BIG5 = indicator variable that equals 1 if audited by a Big 5 firm, and 0 otherwise;
LASSETS = natural log of total assets;
SECURITIES = [1–(total securities/total assets)];
SNPL = non-performing loans divided by lagged total loans;
LOSS = indicator variable that equals 1 if net income < 0, and 0 otherwise;
INTANG = intangible assets divided by total assets;

[17] As in Fields et al. (2004), this model includes three loan categories. We also estimate unexpected fees after including all the six loan categories used in model (1). The relations between *ALLP* and *UFEE* are qualitatively the same when we use all six loan categories to estimate abnormal fees.

EFFICIENCY = total operating expenses divided by total revenues (measured as net interest income plus non-interest income);
SLCO = net loan charge-offs divided by loan loss allowance;
SCOMM = total commercial and agricultural loans divided by total loans;
SCON = total consumer loans divided by total loans;
SRESTATE = total real estate loans divided by total loans;
TCAP = total risk-adjusted capital ratio; and
EXEMPT = indicator variable that equals 1 if a bank has less than $500 million in total assets during the years 2000 through 2003 or less than $500 million in total assets and public float less than $75 million for the year 2004 or less than $1 billion in total assets and public float less than $75 million for the years 2005 and 2006, and 0 otherwise.

We use the residuals from model (2) as the unexpected (abnormal) fees for our four fee measures.

Main Tests

Our objective is to test whether the association between abnormal LLP (*ALLP*) and fees paid to the auditor is stronger for banks that are exempt from the annual auditing and reporting of internal control assessments under FDICIA and Section 404 of SOX. We create *EXEMPT*, an indicator variable that equals 1 for banks that are exempt from FDICIA and Section 404 of SOX and 0 otherwise, and interact *EXEMPT* with fees paid to the auditor. This specification allows us to test whether the effect of fee dependence on earnings management is more pronounced for small banks that are subject to less scrutiny than large banks.

We estimate the model separately for negative (income-increasing) *ALLP* and positive (income-decreasing) *ALLP*.[18] Negative *ALLP* are of particular interest because of their positive impact on reported earnings. We control for the following factors that prior research has documented are associated with abnormal accruals (Ashbaugh et al. 2003): firm size, auditor type, market-to-book ratio, level of accruals, and performance. We use log of market value of equity to measure size, and represent performance by two variables, existence of loss and earnings before LLP, and growth by market-to-book ratio. We control for the reversal of accruals over time by including past LLP, and for capital management incentives by including beginning of year tier 1 capital ratio and total capital ratio.[19] We estimate the following model:

$$\begin{aligned} ALLP = {} & \gamma_0 + \gamma_1 EXEMPT + \gamma_2 UFEE + \gamma_3 EXEMPT \times UFEE + \gamma_4 BIG5 + \gamma_5 MB + \gamma_6 LMVE \\ & + \gamma_7 LOSS + \gamma_8 PASTLLP + \gamma_9 EBP + \gamma_{10} TIER1_{t-1} + \gamma_{11} TCAP_{t-1} + YEARCONTROLS \\ & + \varepsilon \end{aligned} \tag{3}$$

where:

ALLP = abnormal loan loss provision (the residual from model (1));

[18] We report the results of estimating model (3) using both positive and negative *ALLP* in Table 6.

[19] Banks had incentives to influence regulatory capital through LLP prior to the change in bank capital adequacy requirements in 1990. This change altered banks' incentives to manage capital through LLP as loan loss allowance is no longer considered part of tier I or core capital. Therefore, we do not expect γ_{10} to be significantly different from 0. Furthermore, loan loss allowance is included in tier II or supplementary capital only up to 1.25 percent of risk-adjusted assets. Therefore, we predict a negative coefficient on γ_{10} as higher capital ratios would require lower LLP, *ceteris paribus*.

EXEMPT= indicator variable that equals 1 if a bank has less than $500 million in total assets during the years 2000 through 2003 or less than $500 million in total assets and public float less than $75 million for the year 2004 or less than $1 billion in total assets and public float less than $75 million for the years 2005 and 2006, and 0 otherwise;
UFEE = unexpected total fees or nonaudit fees or audit fees or fee ratio from model (2);
BIG5 = indicator variable that equals 1 if the bank is audited by a Big 5 firm, and 0 otherwise;
MB = market-to-book ratio at the end of the year;
LMVE = natural log of market value of common equity at the end of the year;
LOSS = indicator variable that equals 1 if net income < 0, and 0 otherwise;
PASTLLP= prior year's LLP divided by total assets at the beginning of the year;
EBP = net income before extraordinary items and loan loss provisions divided by total assets at the beginning of the year;
TIER1 = tier 1 risk-adjusted capital ratio at the beginning of the year; and
TCAP = total risk-adjusted capital ratio at the beginning of the year.

The coefficients of interest in model (3) are γ_2 and γ_3. For negative (income-increasing) *ALLP*, a negative γ_2 is consistent with auditor fee dependence on the client, and indicates that higher fees are associated with greater (more negative) income-increasing *ALLP* for large banks that are not exempt from regulation. A negative γ_3 indicates greater fee dependence on clients that are exempt from regulation relative to large banks that are not. For positive (income-decreasing) *ALLP*, a negative γ_2 implies that higher auditor fees are associated with less conservative earnings management (i.e., income-decreasing *ALLP* are less positive) for large banks that are not exempt from regulation. Similarly, a negative γ_3 indicates greater fee dependence on clients that are exempt from regulation relative to large banks that are not.

Kinney and Libby (2002) indicate that both audit fees and nonaudit fees are capable of increasing the economic bond. They reason that the auditor might be willing to decrease audit fees or nonaudit fees to maximize potential total fees. Therefore, it is the total fees derived from a single client, rather than just audit fees, that best reflect the economic bond. Prior research also employs nonaudit fees to measure auditor independence because lower regulation and the wide variety of nonaudit services (from information technology to tax preparation) allow for inclusion of rents in nonaudit fees (Srinidhi and Gul 2007). Kinney and Libby (2002, 109) state, "more insidious effects on economic bond may result from *unexpected* nonaudit and audit fees that may more accurately be likened to attempted bribes" (emphasis in the original). This suggests that unexpected fees are a better measure of the auditor-client economic bond because they reflect the "excess" profit derived from the audit client. Another benefit of using unexpected fees is that we explicitly control for several risk-related variables that we include in the estimation of expected fees. Therefore, we use unexpected total fees and unexpected nonaudit fees as our primary measures of the economic bond. However, for consistency with prior research (e.g., Ashbaugh et al. 2003; Srinidhi and Gul 2007) we also conduct our analysis with unexpected audit fees and unexpected fee ratio (nonaudit fees/total fees) as alternate measures of the economic bond. Although we use unexpected fees as our primary measure of economic bonding, we also estimate a variation of model (3) where we replace *UFEE* with actual fees paid to the auditor. We discuss sample selection in the next section.

IV. SAMPLE SELECTION

We select sample banks from banks listed in the 2007 Bank Compustat annual data files and obtain auditor fees for the years 2000–2006 from Audit Analytics. The intersection of the Audit Analytics with Bank Compustat data results in an initial sample of 2,044 bank-year observations.

We hand-collect non-performing loans data for the period 1999–2006 from annual reports and obtain data on loan portfolio composition from the Federal Reserve Bank Holding Company Database (call reports). Our final sample with available data for all variables and without outliers comprises 1,740 bank-year observations for 301 banks.[20] We have 1,618 observations with available data for the models that use abnormal fee measures. The sample does not include any privately held banks.

Table 2, Panel A reports descriptive statistics for the scaled variables used in the regressions. On average, nonaudit fees as a proportion of total fees (*FEERATIO*) is about 0.32. The mean values of *UTOTFEE*, *UNAFEE*, and *ALLP* are zero by construction. Note that *UTOTFEE* and *UNAFEE* are estimated from 1,618 observations. More than 52 percent of the sample observations are audited by Big 5 auditors. Turning to the bank loan variables, the ratios of average *LLP*, loan charge-offs, and beginning non-performing loans to beginning total assets are 0.003, 0.002, and 0.004, respectively.[21] About 19.2 percent of total observations are exempt from Section 36 of the FDI Act and Section 404 of SOX.[22]

Table 2, Panel B reports correlations for the scaled dependent and independent variables. As expected, LLP is positively correlated with non-performing loans (*BEGNPL*), change in non-performing loans (*CHNPL*), and loan charge-offs (*LCO*). *LLP* is also positively correlated with total loans (*LOANS*) and change in loans outstanding (*CHLOANS*). LLP is significantly ($p < 0.05$) positively related to log of total fees and log of nonaudit fees, and *ALLP* is weakly negatively correlated with nonaudit fees.

V. RESULTS

Estimation of Abnormal LLP

We report the estimation results of model (1) in Table 3, Panel A. The t-statistics reported in Table 3 and in the other tables are based on standard errors adjusted for firm-level clustering. The results in Table 3 show that the coefficients on the determinants of LLP have the expected signs except for the coefficient on *BEGNPL*, which is not significant. The coefficients on *BEGLLA*, *CHNPL*, *LCO*, and *LOANS* are significant at the 0.01 level. Of the variables that reflect loan type, only the coefficients on commercial and consumer loans are statistically significant. The explanatory power of the model is high (adjusted $R^2 = 66.32$ percent), indicating that the model describes the variation in LLP quite well.[23]

Estimation of Abnormal Fees

Table 3, Panel B reports the results from estimating model (2). While we report results of model (2) for four fee measures—total fees, nonaudit fees, audit fees, and fee ratio—our primary focus is on total fees and nonaudit fees. As expected, we find a positive relation between the fee

[20] We delete observations in the top 1 percent of scaled loan loss provisions, loan charge-offs, beginning non-performing loans, total loans, and market-to-book ratios. The values of all variables in the reduced sample are generally within three standard deviations from the mean.

[21] These ratios are lower than the values reported in Wahlen (1994) and Kanagaretnam et al. (2004) due to increase in overall quality of loans (Edwards and Mishkin 1995). Similarly, our sample data indicate that mean LLP equals 17 percent of earnings before provisions, considerably lower than the 148 percent for Wahlen's (1994) sample and lower than the 31 percent for Kanagaretnam et al.'s (2004) sample.

[22] We closely follow the approach of Gao et al. (2009) for identifying accelerated/non-accelerated filers under Section 404 of SOX. Our sample includes 14 bank-year observations that have assets less than $500 million in 2004 but were accelerated filers. For 2005 and 2006, it includes 34 bank-year observations that had assets between $500 million and $1 billion that were non-accelerated filers.

[23] Even without the year controls, the adjusted R^2 is 56.55 percent, indicating that our model fit compares favorably with the results of prior research. For example, McNichols (2002) reports that when change in working capital accruals is regressed on changes in sales and property, plant, and equipment, the adjusted R^2 for the pooled sample is 7.3 percent. Ashbaugh et al. (2003) report mean R^2 values of 11 percent and 18 percent for their models of abnormal accruals.

TABLE 2

Descriptive Statistics and Correlations

Panel A: Main Variables Used in Regressions (n = 1740)

Variable	Mean	Std. Dev.	Minimum	25%	Median	75%	Maximum
LTOTFEE	12.6647	1.1829	10.5348	11.8373	12.5183	13.2761	16.4232
LNAFEE	11.2298	1.5040	7.8478	10.3008	11.0895	12.0438	15.9673
LAFEE	12.2233	1.1668	9.9034	11.4081	12.0857	12.8605	15.7665
FEERATIO	0.3151	0.2037	0.0105	0.1605	0.2688	0.4414	0.8413
UTOTFEE	0.0000	0.5142	−1.3148	−0.3025	−0.0027	0.3136	1.1758
UNAFEE	0.0000	1.0379	−3.4709	−0.4973	0.0533	0.6302	2.1350
LLP	0.0025	0.0021	−0.0009	0.0011	0.0021	0.0033	0.0104
ALLP	0.0000	0.0012	−0.0029	−0.0006	−0.0001	0.0001	0.0039
LOANS	0.7452	0.1764	0.3324	0.6382	0.7285	0.8453	1.2220
BEGLLA	0.0090	0.0030	0.0031	0.0074	0.0088	0.0122	0.0179
CHLOANS	0.0984	0.1101	−0.0766	0.0336	0.0728	0.1338	0.4166
BEGNPL	0.0035	0.0032	0.0000	0.0012	0.0028	0.0049	0.0144
CHNPL	0.0005	0.0029	−0.0062	−0.0001	0.0000	0.0012	0.0101
LCO	0.0017	0.0017	−0.0003	0.0005	0.0013	0.0023	0.0080
EXEMPT	0.1920	0.3941	0	0	0	0	1
BIG5	0.5247	0.4995	0	0	1	1	1
MB	2.9110	1.7626	0.9404	1.8040	2.3690	3.4230	9.9947
LMVE	19.7910	1.5186	17.4361	18.7214	19.5704	20.5792	24.8273
LOSS	0.0075	0.0861	0	0	0	0	1
PASTLLP	0.0024	0.0024	−0.0005	0.0012	0.0020	0.0030	0.0103
EBP	0.0146	0.0053	−0.0565	0.0118	0.0144	0.0173	0.0279
TIER1	11.9085	3.4947	7.1700	9.9500	11.2150	12.9800	20.1200
TCAP	13.5740	3.3545	9.780	11.630	12.8000	14.5450	21.8600

Panel B: Correlations between Dependent and Independent Variables (n = 1740)

	ALLP	*LOANS*	*BEGLLA*	*CHLOANS*	*BEGNPL*	*CHNPL*	*LCO*	*LTOTFEE*	*LNAFEE*
LLP	0.577	0.234	0.172	0.125	0.230	0.174	0.738	0.059	0.129
	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.013	0.001

(continued on next page)

Panel B: Correlations between Dependent and Independent Variables (n = 1740)

	ALLP	*LOANS*	*BEGLLA*	*CHLOANS*	*BEGNPL*	*CHNPL*	*LCO*	*LTOTFEE*	*LNAFEE*
ALLP		0.000	0.000	0.000	0.000	0.000	0.000	−0.033	−0.040
		1.000	1.000	1.000	1.000	1.000	1.000	0.167	0.098
LOANS		1.000	0.256	0.759	0.016	0.109	−0.039	−0.173	−0.115
			0.001	0.001	0.497	0.001	0.104	0.001	0.001
BEGLLA			1.000	−0.014	0.312	−0.062	0.327	0.033	0.068
				0.549	0.001	0.010	0.001	0.167	0.005
CHLOANS				1.000	−0.125	0.095	−0.159	−0.081	−0.064
					0.001	0.001	0.001	0.001	0.008
BEGNPL					1.000	−0.203	0.379	0.014	0.018
						0.001	0.001	0.560	0.442
CHNPL						1.000	0.074	−0.046	−0.023
							0.002	0.057	0.335
LCO							1.000	0.206	0.230
								0.001	0.001
LTOTFEE								1.000	0.807
									0.001
LNAFEE									1.000

	LAFEE	*FEERATIO*	*EXEMPT*	*BIG5*	*MB*	*LMVE*	*LOSS*	*PASTLLP*	*EBP*	*TCAP*
LLP	−0.004	0.187	−0.039	0.044	0.132	0.052	−0.002	0.461	0.298	−0.167
	0.874	0.001	0.100	0.070	0.001	0.030	0.935	0.001	0.001	0.001
ALLP	−0.038	−0.003	−0.054	−0.026	0.067	−0.035	0.013	0.121	0.141	−0.021
	0.112	0.911	0.058	0.277	0.005	0.141	0.577	0.001	0.001	0.386
LOANS	−0.187	0.034	0.100	−0.211	0.167	−0.161	−0.033	0.025	0.193	−0.282
	0.001	0.154	0.001	0.001	0.001	0.001	0.169	0.301	0.001	0.001
BEGLLA	0.009	0.079	−0.020	−0.009	0.024	0.004	0.119	0.471	0.115	−0.125
	0.697	0.001	0.402	0.697	0.324	0.854	0.001	0.001	0.001	0.001
CHLOANS	−0.081	−0.004	0.150	−0.121	0.204	−0.051	−0.038	−0.092	0.158	−0.040
	0.001	0.852	0.001	0.001	0.001	0.033	0.112	0.001	0.001	0.094
BEGNPL	0.008	0.018	−0.008	0.008	−0.052	−0.048	0.047	0.297	−0.007	−0.070
	0.729	0.465	0.728	0.725	0.029	0.047	0.051	0.001	0.757	0.004

(continued on next page)

	LAFEE	FEERATIO	EXEMPT	BIG5	MB	LMVE	LOSS	PASTLLP	EBP	TCAP
CHNPL	−0.061	0.043	0.031	−0.037	−0.010	−0.056	−0.042	−0.055	0.066	−0.026
	0.011	0.074	0.192	0.122	0.686	0.020	0.078	0.022	0.006	0.285
LCO	0.157	0.154	−0.108	0.131	−0.022	0.152	0.025	0.573	0.193	−0.122
	0.001	0.001	0.001	0.001	0.354	0.001	0.289	0.001	0.001	0.001
LTOTFEE	0.944	0.166	−0.442	0.531	−0.014	0.829	0.020	0.068	0.094	−0.204
	0.001	0.001	0.001	0.001	0.554	0.001	0.416	0.005	0.001	0.001
LNAFEE	0.613	0.635	−0.312	0.418	0.097	0.690	0.005	0.107	0.119	−0.188
	0.001	0.001	0.001	0.001	0.001	0.001	0.839	0.001	0.001	0.001
LAFEE	1.000	−0.158	−0.437	0.498	−0.084	0.777	0.019	0.038	0.061	−0.175
		0.001	0.001	0.001	0.000	0.001	0.441	0.117	0.011	0.001
FEERATIO		1.000	−0.013	0.098	0.218	0.168	0.006	0.101	0.104	−0.086
			0.576	0.001	0.001	0.001	0.790	0.001	0.001	0.000
EXEMPT			1.000	−0.367	0.016	−0.487	0.038	−0.067	−0.170	0.205
				0.001	0.502	0.001	0.115	0.005	0.001	0.001
BIG5				1.000	0.167	0.532	0.029	0.045	0.082	−0.149
					0.001	0.001	0.225	0.061	0.001	0.001
MB					1.000	0.268	0.050	0.062	0.184	−0.205
						0.001	0.036	0.009	0.001	0.001
LMVE						1.000	−0.060	0.033	0.288	−0.207
							0.012	0.172	0.001	0.001
LOSS							1.000	0.113	−0.415	0.002
								0.001	0.001	0.929
PASTLLP								1.000	0.062	−0.119
									0.009	0.001
EBP									1.000	0.074
										0.002

n = 1618 for *UTOTFEE* and *UNAFEE*.
Please see Table 1 for variable definitions.
Numbers below the correlation coefficient are p-values.

TABLE 3

Regressions for Computing *ALLP* and Abnormal Fees

Panel A: Results of Regression of *LLP* on Determinants of Normal *LLP*[a]

Variable	Expected Sign	Coefficient	t-statistic
Intercept	?	−0.0012***	−5.87
BEGLLA	−	−0.1148***	−4.66
BEGNPL	+	−0.0072	−0.43
CHNPL	+	0.0406***	2.69
LCO	+	0.9559***	25.88
CHLOANS	?	0.0006	0.69
LOANS	+	0.0034***	7.66
COMM	?	0.0004***	2.89
CON	?	−0.0014***	−2.76
RESTATE	?	0.0000	0.10
AGRI	?	0.0007	0.34
FBG	?	−0.0209	−1.39
DEPINS	?	0.0072	0.73
Year Controls		Yes	
n		1740	
Adjusted R^2		66.32%	

Panel B: Results of Regressions of Fee Measures on Determinants of Fees[b]

Variable	Total Fee	Nonaudit Fee	Audit Fee	Fee Ratio
(1)	(2)	(3)	(4)	(5)
Intercept	5.4366***	2.7859***	5.7552***	−0.0626
	(12.58)	(3.34)	(12.47)	(−0.44)
BIG5	0.2908***	0.2693***	0.2508***	0.0193
	(5.63)	(2.77)	(4.87)	(1.13)
LASSETS	0.6714***	0.7136***	0.6292***	0.0188**
	(22.53)	(13.71)	(21.50)	(2.44)
SECURITIES	0.8548***	1.1918***	0.7794***	0.0678
	(4.00)	(3.49)	(3.10)	(1.11)
SNPL	13.9441***	41.8254***	1.5396	6.6917***
	(2.61)	(4.56)	(0.35)	(4.76)
LOSS	0.2522	−0.3621	0.4643*	−0.0950*
	(1.08)	(−0.79)	(1.90)	(−1.35)
INTANG	5.4330***	−1.0942	8.4852***	−1.5201***
	(3.92)	(−0.38)	(4.83)	(−3.43)
EFFICIENCY	1.1689***	1.7519***	0.7808***	0.1665*
	(3.92)	(2.94)	(2.76)	(1.80)
SLCO	−0.0863	−0.3179	0.0303	−0.0718*
	(−0.85)	(−1.46)	(0.24)	(−1.68)
SCOMM	0.2634***	0.2556	0.2497**	0.0021
	(2.68)	(1.55)	(2.22)	(0.06)
SCON	−0.4945**	−0.2637	−0.5205**	0.0339
	(−2.15)	(−0.60)	(−2.06)	(0.39)

(continued on next page)

Panel B: Results of Regressions of Fee Measures on Determinants of Fees[b]

Variable	Total Fee	Nonaudit Fee	Audit Fee	Fee Ratio
(1)	(2)	(3)	(4)	(5)
SRESTATE	−0.0812	−0.1646	−0.0511	−0.0181
	(−1.20)	(−1.20)	(−0.80)	(−0.68)
TCAP	0.0159***	0.0061	0.0223***	−0.0027
	(2.77)	(0.35)	(3.19)	(−0.74)
EXEMPT	−0.0167	−0.1237	−0.0589	0.0281
	(−0.28)	(−1.26)	(−0.94)	(1.57)
Year controls	Yes	Yes	Yes	Yes
n	1618	1618	1618	1618
Adj. R^2	80.38%	51.11%	74.44%	6.53%

*, **, *** Indicate 0.10, 0.05, and 0.01 significance levels, respectively, for a two-tailed test.

[a] Panel A: The regression model is:

$$LLP = \alpha_0 + \alpha_1 BEGLLA + \alpha_2 BEGNPL + \alpha_3 CHNPL + \alpha_4 LCO + \alpha_5 CHLOANS + \alpha_6 LOANS + LOANCATEGORIES + YEARCONTROLS + \varepsilon.$$

t-statistics are based on standard errors adjusted for firm-level clustering.

[b] Panel B: The regression model is:

$$FEE = \delta_0 + \delta_1 BIG5 + \delta_2 LASSETS + \delta_3 SECURITIES + \delta_4 SNPL + \delta_5 LOSS + \delta_6 INTANG + \delta_7 EFFICIENCY + \delta_8 SLCO + \delta_9 SCOMM + \delta_{10} SCON + \delta_{11} SCESTATE + \delta_{12} TCAP + \delta_{13} EXEMPT + YEARCONTROLS + \varepsilon.$$

t-statistics, in parentheses, are based on standard errors adjusted for firm-level clustering.
Please see Table 1 for variable definitions.

measures, total fees and nonaudit fees, and *BIG5*, *LASSETS*, *SECURITIES*, *NPL*, *EFFICIENCY*, and *SCOMM*. The signs of the coefficients are generally consistent with Fields et al. (2004). For the total fees model, the adjusted R^2 is 80.38 percent, indicating a good fit. For the nonaudit fees model, the adjusted R^2 is 51.11 percent. These values are higher than the adjusted R^2 values reported in Ashbaugh et al. (2003) for industrial firms and consistent with the R^2 values reported in Fields et al. (2004). We use the residuals from model (2) as the unexpected (abnormal) fees for our fee measures.[24]

Association between Income-increasing (Negative) Abnormal LLP and Abnormal and Actual Fee Measures

We present the results of model (3) relating abnormal LLP to fee measures (total fees, nonaudit fees, audit fees, and fee ratio) in Tables 4 and 5. Although we report estimation results for all four fee measures for completeness, we only discuss the results for unexpected total and nonaudit fees, the measures of particular interest.[25] Table 4 reports the results for negative (income-increasing) *ALLP*, and Table 5 reports the results for positive (income-decreasing) *ALLP*. In each table, Panel A presents the results using abnormal (unexpected) fee measures, and Panel B presents the results using actual fee measures.

[24] The correlations between our proxies for abnormal LLP and abnormal fee measures and size are very low, confirming that our subsequent analysis for large and small banks does not suffer from latent size issues.

[25] The results for audit fees and fee ratio are discussed in the section on sensitivity analysis.

TABLE 4

Relation Between Income-Increasing (Negative) *ALLP* and Fee Measures

Panel A: Results of Regressions of Income-Increasing (Negative) *ALLP* on Abnormal Fee Measures[a]

Variable	Abnormal Total Fee		Abnormal Nonaudit Fee		Abnormal Audit Fee		Abnormal Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	−0.0015**	−0.0015**	−0.0016**	−0.0017**	−0.0016**	−0.0016**	−0.0017**	−0.0018***
	(−2.28)	(−2.20)	(−2.40)	(−2.45)	(−2.36)	(−2.39)	(−2.51)	(−2.64)
EXEMPT	−0.0002**	−0.0002**	−0.0002**	−0.0001**	−0.0002**	−0.0002**	−0.0002**	−0.0002**
	(−2.36)	(−2.27)	(−2.39)	(−2.21)	(−2.35)	(−2.36)	(−2.43)	(−2.33)
UTOTFEE	−0.0001	0.0000						
	(−0.74)	(0.09)						
EXEMPT × *UTOTFEE*		−0.0006***						
		(−2.67)						
UNAFEE			−0.0001*	−0.0000				
			(−1.74)	(−0.88)				
EXEMPT × *UNAFEE*				−0.0003**				
				(−2.42)				
UAFEE					0.0000	0.0000		
					(0.29)	(0.19)		
EXEMPT × *UAFEE*						0.0001		
						(0.40)		
UFEERATIO							−0.0003*	−0.0002
							(−1.78)	(−0.75)
EXEMPT × *UFEERATIO*								−0.0014***
								(−2.81)
BIG5	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0001	0.0000
	(0.41)	(0.51)	(0.25)	(0.39)	(0.36)	(0.35)	(0.63)	(0.26)
MB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.70)	(0.89)	(0.65)	(0.76)	(0.70)	(0.68)	(0.63)	(0.70)
LMVE	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
	(1.17)	(1.05)	(1.30)	(1.25)	(1.25)	(1.27)	(1.39)	(1.41)

(continued on next page)

Panel A: Results of Regressions of Income-Increasing (Negative) *ALLP* on Abnormal Fee Measures[a]

Variable	Abnormal Total Fee		Abnormal Nonaudit Fee		Abnormal Audit Fee		Abnormal Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LOSS	−0.0024**	−0.0024**	−0.0024**	−0.0024**	−0.0022**	−0.0024**	−0.0024**	−0.0024**
	(−2.37)	(−2.44)	(−2.23)	(−2.14)	(−2.17)	(−2.19)	(−2.17)	(−2.00)
PASTLLP	−0.0288	−0.0285*	−0.0300*	−0.0291*	−0.0291*	−0.0290*	−0.0306*	−0.0291*
	(−1.64)	(−1.67)	(−1.75)	(−1.73)	(−1.69)	(−1.69)	(−1.79)	(−1.70)
EBP	−0.0143	−0.0124	−0.0147	−0.0131	−0.0124	−0.0140	−0.0142	−0.0124
	(−1.49)	(−1.29)	(−1.52)	(−1.36)	(−1.45)	(−1.45)	(−1.47)	(−1.29)
TIER1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.82)	(0.54)	(0.76)	(0.64)	(1.27)	(0.90)	(0.85)	(0.81)
TCAP	−0.0000	0.0000	−0.0000	0.0000	−0.0000	−0.0000	−0.0000	−0.0000
	(−0.15)	(0.10)	(−0.14)	(0.03)	(−0.78)	(−0.22)	(−0.22)	(−0.08)
Year Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	882	882	882	882	882	882	882	882
Adj. R^2	5.69%	6.34%	5.95%	6.48%	5.64%	5.55%	5.97%	6.70%

Panel B: Results of Regressions of Income-Increasing (Negative) *ALLP* on Actual Fee Measures[b]

Variable	Total Fee		Nonaudit Fee		Audit Fee		Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	−0.0027***	−0.0026**	−0.0023***	−0.0022***	−0.0024***	−0.0023***	−0.0016***	−0.0016***
	(−2.64)	(−2.50)	(−2.89)	(−2.79)	(−3.08)	(−2.94)	(−2.79)	(−2.77)
EXEMPT	−0.0003	−0.0013	−0.0003	−0.0006	−0.0003	−0.0012	−0.0003	−0.0004
	(−1.06)	(−1.50)	(−0.97)	(−1.01)	(−1.10)	(−1.52)	(−1.04)	(−1.00)
LTOTFEE	−0.0001*	−0.0001						
	(−1.76)	(−0.95)						
EXEMPT × LTOTFEE		−0.0009**						
		(−2.55)						
LNAFEE			−0.0001*	−0.0000				
			(−1.87)	(−0.92)				

(continued on next page)

Panel B: Results of Regressions of Income-Increasing (Negative) *ALLP* on Actual Fee Measures[b]

Variable	Total Fee		Nonaudit Fee		Audit Fee		Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EXEMPT × LNAFEE				−0.0004**				
				(−2.42)				
LAFEE					−0.0001	−0.0001		
					(−1.63)	(−1.48)		
EXEMPT × LAFEE						−0.0003		
						(−1.56)		
FEERATIO							−0.0001	−0.0001
							(−0.68)	(−0.57)
EXEMPT × FEERATIO								−0.0018
								(−0.51)
BIG5	0.0000	0.0000	−0.0000	0.0000	0.0000	0.0000	−0.0000	−0.0000
	(0.35)	(0.37)	(−0.02)	(0.01)	(0.29)	(0.36)	(−0.12)	(−0.12)
MB	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
	(−0.73)	(−0.70)	(−0.65)	(−0.64)	(−0.70)	(−0.68)	(−0.52)	(−0.52)
LMVE	0.0001*	0.0002*	0.0001**	0.0001*	0.0001**	0.0001**	0.0001	0.0001
	(1.94)	(1.82)	(1.95)	(1.87)	(2.18)	(2.07)	(1.58)	(1.57)
LOSS	0.0005*	0.0004*	0.0004*	0.0004**	0.0005*	0.0004**	0.0004*	0.0004*
	(1.80)	(1.75)	(1.71)	(1.98)	(1.82)	(2.02)	(1.68)	(1.68)
PASTLLP	−0.0529*	−0.0530*	−0.0534*	−0.0530*	−0.0535*	−0.0540*	−0.0545*	−0.0542*
	(−1.83)	(−1.83)	(−1.84)	(−1.83)	(−1.84)	(−1.86)	(−1.88)	(−1.86)
EBP	−0.0073	−0.0074	−0.0067	−0.0065	−0.0057	−0.0061	−0.0044	−0.0043
	(−0.64)	(−0.65)	(−0.61)	(−0.59)	(−0.53)	(−0.58)	(−0.41)	(−0.39)
TIER1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.06)	(0.07)	(0.10)	(0.11)	(0.11)	(0.09)	(0.28)	(0.29)
TCAP	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
	(−0.46)	(−0.45)	(−0.53)	(−0.54)	(−0.48)	(−0.40)	(−0.61)	(−0.62)
Year Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	952	952	952	952	952	952	952	952
Adj. R^2	5.77%	6.68%	5.81%	6.12%	5.61%	6.34%	5.42%	5.59%

(continued on next page)

*, **, *** Indicate 0.10, 0.05, and 0.01 significance levels, respectively, for a two-tailed test. t-statistics, in parentheses, are based on standard errors adjusted for firm-level clustering.

[a] Panel A: The regression model is:

$$ALLP = \gamma_0 + \gamma_1 EXEMPT + \gamma_2 UFEE + \gamma_3 EXEMPT \times UFEE + \gamma_4 BIG5 + \gamma_5 MB + \gamma_6 LMVE + \gamma_7 LOSS + \gamma_8 PASTLLP + \gamma_9 EBP + \gamma_{10} TIER1_{t-1} + \gamma_{11} TCAP_{t-1} + YEARCONTROLS + \varepsilon.$$

[b] Panel B: The regression model is:

$$ALLP = \gamma_0 + \gamma_1 EXEMPT + \gamma_2 FEE + \gamma_3 EXEMPT \times FEE + \gamma_4 BIG5 + \gamma_5 MB + \gamma_6 LMVE + \gamma_7 LOSS + \gamma_8 PASTLLP + \gamma_9 EBP + \gamma_{10} TIER1_{t-1} + \gamma_{11} TCAP_{t-1} + YEARCONTROLS + \varepsilon.$$

Please see Table 1 for variable definitions.

For each fee measure, we present results for two specifications. The first specification estimates the relation between *ALLP* and each abnormal and actual fee measure for all banks. The second specification (column 3) distinguishes between small banks that are exempt from FDICIA and Section 404 of SOX and large banks, thereby allowing the relation between *ALLP* and fee measures to vary across these two classes of banks.[26]

The results in column 2 of Table 4, Panel A indicate that the coefficient on *UTOTFEE* is not significant, but the results in column 4 indicate that the coefficient on *UNAFEE* is significant at the 0.10 level. These results suggest that income-increasing earnings management via *ALLP* is increasing in unexpected nonaudit fees. In terms of the control variables, *LOSS* and *PASTLLP* (in column 4) are significantly related to *ALLP*. Interestingly, there is no difference in abnormal LLP between clients of *BIG5* and non-*BIG5* auditors. As expected, the capital ratios are not significantly associated with *ALLP*, confirming the reduced capital management incentives through *ALLP* in the post-1990 period.

Turning to the second specification (columns 3 and 5 of Table 4), the coefficients on *UTOTFEE* and *UNAFEE* are not significant, indicating no relation between income-increasing earnings management and unexpected fees for large banks.[27] In contrast, the coefficients on *EXEMPT* $\times$ *UTOTFEE* and *EXEMPT* $\times$ *UNAFEE* are negative and significant at the 0.05 level. In addition, both the sum of the coefficients on *UTOTFEE* and *EXEMPT* $\times$ *UTOTFEE* and on *UNAFEE* and *EXEMPT* $\times$ *UNAFEE* are significantly negative at the 0.01 level. These results indicate that income-increasing *ALLP* is higher for audit clients that are small banks (and, therefore, are exempt from the regulatory oversight) that pay higher abnormal total fees or abnormal nonaudit fees to their auditors.

The results in Table 4, Panel B indicate that the coefficients on *EXEMPT* $\times$ *LTOTFEE* and *EXEMPT* $\times$ *LNAFEE* are negative and significant at the 0.05 level.[28] Overall, these results are consistent with the results based on unexpected fee measures reported in Panel A, indicating that our findings are not sensitive to the use of actual or unexpected fees. In addition, the results document that only small banks that are exempt from regulatory compliance exhibit a significant relation between earnings management via income-increasing LLP and fees paid to the auditor. In short, our findings suggest that the potential impairment of auditor independence may be serious only for small banks and not for large banks.

Association between Income-Decreasing (Positive) Abnormal LLP and Abnormal and Actual Fee Measures

The results reported in Table 5, Panel A show that *UTOTFEE* is not significantly associated with income-decreasing *ALLP*. However, the coefficients on *EXEMPT* $\times$ *UTOTFEE* and *EXEMPT* $\times$ *UNAFEE* are negative and significant at the 0.05 level.[29] These results indicate that higher unexpected total fees and unexpected nonaudit fees paid to auditors of exempt small banks are associated with a lower level of income-decreasing *ALLP*. Turning to columns 2 and 3 of Panel B, *LTOTFEE* is positively and significantly associated with *ALLP* at the 0.01 level in each of those regressions. One interpretation of this result is that clients of auditors who pay higher total fees

[26] For the abnormal fee regressions, we have 170 (70) bank-year (banks) observations for which *EXEMPT* = 1. For the actual fee regressions, we have 177 (72) bank-year (banks) observations for which *EXEMPT* = 1.

[27] We also estimate model (3) using only observations for which *EXEMPT* = 0 (i.e., large banks). We find that the coefficients on *UTOTFEE* and *UNAFEE* are insignificant at the 0.20 level for both income-increasing and income-decreasing *ALLP*, indicating no evidence of fee dependence for large banks.

[28] We mean-center the fee variables to reduce problems with multicollinearity among the interaction terms (Neter et al.1989; Aiken and West 1991).

[29] For the abnormal fee regressions, we have 142 (70) bank-year (banks) observations for which *EXEMPT* = 1. For the actual fee regressions, we have 157 (74) bank-year (banks) observations for which *EXEMPT* = 1.

TABLE 5

Relation Between Income-Decreasing (Positive) *ALLP* and Fee Measures

Panel A: Results of Regressions of Income-Decreasing (Positive) *ALLP* on Abnormal Fee Measures[a]

Variable	Abnormal Total Fee		Abnormal Nonaudit Fee		Abnormal Audit Fee		Abnormal Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.0028***	0.0023***	0.0022***	0.0022***	0.0023***	0.0023***	0.0022***	0.0022***
	(3.53)	(3.52)	(3.31)	(3.27)	(3.51)	(3.49)	(3.29)	(3.29)
EXEMPT	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	(0.78)	(0.78)	(0.83)	(0.69)	(0.80)	(0.81)	(0.89)	(0.84)
UTOTFEE	0.0000	0.0001						
	(0.14)	(1.05)						
EXEMPT × UTOTFEE		−0.0003**						
		(−1.96)						
UNAFEE			−0.0001*	−0.0000				
			(−1.82)	(−1.00)				
EXEMPT × UNAFEE				−0.0002**				
				(−2.48)				
UAFEE					0.0000	0.0001		
					(0.43)	(0.80)		
EXEMPT × UAFEE						−0.0001		
						(−0.91)		
UFEERATIO							−0.0003	−0.0002
							(−1.57)	(−0.98)
REG × UFEERATIO								−0.0006
								(−1.21)
BIG5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.43)	(0.49)	(0.35)	(0.54)	(0.38)	(0.37)	(0.26)	(0.36)
MB	0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
	(0.05)	(−1.05)	(−1.15)	(−1.09)	(−1.21)	(−1.04)	(−1.10)	(−1.05)
LMVE	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−0.0001***	−0.0001***
	(−3.13)	(−3.14)	(−2.97)	(−3.00)	(−3.12)	(−3.10)	(−2.97)	(−3.00)

(continued on next page)

Panel A: Results of Regressions of Income-Decreasing (Positive) *ALLP* on Abnormal Fee Measures[a]

Variable	Abnormal Total Fee		Abnormal Nonaudit Fee		Abnormal Audit Fee		Abnormal Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LOSS	0.0013**	0.0011**	0.0013**	0.0015**	0.0013**	0.0013**	0.0010**	0.0014**
	(2.10)	(2.03)	(1.99)	(2.09)	(2.33)	(2.38)	(2.09)	(2.07)
PASTLLP	0.1052***	0.1033***	0.1030***	0.1027***	0.1046***	0.1035***	0.1021***	0.1028***
	(3.09)	(2.64)	(3.65)	(3.70)	(3.25)	(2.91)	(3.59)	(3.55)
EBP	0.0470***	0.0483***	0.0482***	0.0500***	0.0473***	0.0477***	0.0490***	0.0494***
	(2.89)	(2.97)	(2.87)	(2.86)	(2.84)	(2.95)	(2.88)	(2.90)
TIER1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.12)	(0.10)	(0.24)	(0.30)	(0.10)	(0.07)	(0.16)	(0.20)
TCAP	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
	(−0.56)	(−0.55)	(−0.67)	(−0.67)	(−0.54)	(−0.54)	(−0.62)	(−0.62)
Year Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	736	736	736	736	736	736	736	736
Adj. R^2	13.64%	13.99%	14.18%	14.80%	13.66%	13.64%	13.98%	14.03%

Panel B: Results of Regressions of Income-Decreasing (Positive) *ALLP* on Actual Fee Measures[b]

Variable	Total Fee		Nonaudit Fee		Audit Fee		Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.0055***	0.0055***	0.0030***	0.0034***	0.0044***	0.0044***	0.0027***	0.0027***
	(4.17)	(4.18)	(4.25)	(4.26)	(4.20)	(4.20)	(4.30)	(4.29)
EXEMPT	−0.0001	−0.0002	−0.0001	−0.0002	−0.0001	−0.0000	−0.0001	−0.0002
	(−0.68)	(−0.77)	(−0.68)	(−0.91)	(−0.47)	(−0.05)	(−0.60)	(−1.22)
LTOTFEE	0.0002***	0.0002***						
	(3.12)	(3.13)						
EXEMPT × LTOTFEE		−0.0002						
		(−0.57)						
LNAFEE			0.0001	0.0001				
			(1.37)	(1.39)				

(continued on next page)

Panel B: Results of Regressions of Income-Decreasing (Positive) *ALLP* on Actual Fee Measures[b]

Variable	Total Fee		Nonaudit Fee		Audit Fee		Fee Ratio	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EXEMPT × LNAFEE				−0.0001				
				(−0.77)				
LAFEE					0.0002**	0.0002**		
					(2.18)	(2.17)		
EXEMPT × LAFEE						0.0002		
						(0.30)		
FEERATIO							0.0003	0.0003
							(1.07)	(1.11)
EXEMPT × FEERATIO								−0.0013**
								(−2.08)
BIG5	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0000	−0.0000
	(−1.26)	(−1.28)	(−0.55)	(−0.57)	(−1.14)	(−1.14)	(−0.41)	(−0.44)
MB	0.0000	0.0000	0.0000	−0.0000	0.0000	0.0000	−0.0000	−0.0000
	(0.94)	(0.95)	(0.07)	(−0.08)	(0.44)	(0.43)	(−0.20)	(0.20)
LMVE	−0.0003***	−0.0003***	−0.0002***	−0.0002***	−0.0002***	−0.0002***	−0.0001***	−0.0001***
	(−4.99)	(−4.99)	(−3.84)	(−3.83)	(−3.84)	(−3.83)	(−3.49)	(−3.48)
LOSS	0.0021**	0.0021**	0.0022**	0.0021**	0.0022**	0.0021**	0.0021**	0.0021**
	(2.18)	(2.18)	(2.32)	(2.32)	(2.23)	(2.23)	(2.21)	(2.21)
PASTLLP	0.0896***	0.0897***	0.0991***	0.0993***	0.0935***	0.0935***	0.1021***	0.1025***
	(3.03)	(3.03)	(3.35)	(3.36)	(3.13)	(3.13)	(3.50)	(3.50)
EBP	0.0344**	0.0344***	0.0311**	0.0312**	0.0348**	0.0348**	0.0299**	0.0299**
	(2.59)	(2.58)	(2.38)	(2.37)	(2.53)	(2.54)	(2.22)	(2.22)
TIER1	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
	(−0.14)	(−0.14)	(−0.09)	(−0.09)	(−0.13)	(−0.14)	(−0.09)	(−0.09)
TCAP	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
	(−0.22)	(−0.21)	(−0.29)	(−0.28)	(−0.28)	(−0.29)	(−0.28)	(−0.29)
Year Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
n	788	788	788	788	788	788	788	788
Adj. R^2	13.66%	13.68%	12.43%	12.46%	12.74%	12.74%	12.27%	12.32%

(continued on next page)

*, **, *** Indicate 0.10, 0.05, and 0.01 significance levels, respectively, for a two-tailed test.
t-statistics, in parentheses, are based on standard errors adjusted for firm-level clustering.

[a] Panel A: The regression model is:

$$ALLP = \gamma_0 + \gamma_1 EXEMPT + \gamma_2 UFEE + \gamma_3 EXEMPT \times UFEE + \gamma_4 BIG5 + \gamma_5 MB + \gamma_6 LMVE + \gamma_7 LOSS + \gamma_8 PASTLLP + \gamma_9 EBP + \gamma_{10} TIER1_{t-1} + \gamma_{11} TCAP_{t-1} + YEARCONTROLS + \varepsilon.$$

[b] Panel B: The regression model is:

$$ALLP = \gamma_0 + \gamma_1 EXEMPT + \gamma_2 FEE + \gamma_3 EXEMPT \times FEE + \gamma_4 BIG5 + \gamma_5 MB + \gamma_6 LMVE + \gamma_7 LOSS + \gamma_8 PASTLLP + \gamma_9 EBP + \gamma_{10} TIER1_{t-1} + \gamma_{11} TCAP_{t-1} + YEARCONTROLS + \varepsilon.$$

Please see Table 1 for variable definitions.

engage in a higher level of income-decreasing earnings management.[30] Although positive, the insignificant coefficient on *LNAFEE* does not support the total fee results. Further, the coefficients on *EXEMPT* × *LTOTFEE* and *EXEMPT* × *LNAFEE* are insignificant. These results differ from the results in Panel A due to the use of different fee measures. If we do not control for factors determining normal fees, we do not observe a significant relation between income-decreasing *ALLP* and either unexpected total fees or unexpected nonaudit fees (see Panel A), indicating that the results for income-decreasing *ALLP* are sensitive to the use of fees or abnormal fees.

Summary of Results

We briefly summarize our main results on the relation between income-increasing and income-decreasing *ALLP* and unexpected total fees and unexpected nonaudit fees. We find that for large banks, both unexpected total fees and unexpected nonaudit fees have no relationship with either income-increasing or income-decreasing *ALLP*—i.e., unexpected total or nonaudit fees are unrelated to earnings management behavior in large banks. In contrast, we find a significant negative association between income-increasing *ALLP* and both unexpected total fees and unexpected nonaudit fees for small banks relative to large banks. Recall that these small banks are exempt from FDICIA (as well as Section 404 of SOX), which mandated assessments of the effectiveness of internal control over financial reporting by managers and auditors. Furthermore, we also document a significant negative association between unexpected total fees and unexpected nonaudit fees and income-decreasing *ALLP*. These results are consistent with greater earnings management by small banks that pay higher abnormal fees to the auditor. Collectively, our results suggest that auditor independence may be lower when auditor fee dependence on the audit client is higher, but only for small banks that are not subject to the same level of regulatory scrutiny as large banks. We conjecture that auditors of small, exempt banks may not be as concerned with earnings management perhaps because the failure of such a bank, if it occurs, will cause less damage to the auditor's reputation than will the failure of a large bank. Alternatively, partners auditing small, exempt banks may not be as closely monitored internally as partners auditing large banks.[31]

Pooled *ALLP* Regressions

For consistency with prior research on banks (e.g., Ahmed et al. 1999; Kanagaretnam et al. 2004), we also estimate model (3) for the full sample (i.e., combining income-increasing and income-decreasing *ALLP*) and present these results in Table 6. The coefficients on *EXEMPT* × *UTOTFEE* and *EXEMPT* × *UNAFEE* are both negative and significant at the 0.05 level, indicating that our finding of a positive relation between higher abnormal auditor fees and aggressive earnings management for exempted banks is robust.

Association between Abnormal LLP and Abnormal Fee Measures in the Post-SOX Era

While banks with total assets of less than $500 million (less than $1 billion effective 2005) were not subject to the annual reporting requirements mandated by Section 36 of the FDI Act, the passage of SOX (Section 404) also exempts assessment of the effectiveness of internal control over financial reporting by managers and auditors of small banks (banks with public float less than $75 million). Therefore, we expect our results on the association between *ALLP* and unexpected fees observed for small (exempted) banks to be similar in the post-SOX era. We examine whether our results hold for a class of banks that are exempt from both FDICIA and SOX. Section 404 of

[30] Also, note that *EBP* is positive and significant in both panels of Table 5, suggesting that firms are more willing to take income-decreasing *LLP* when they have high earnings. This is consistent with income smoothing.

[31] We thank an anonymous reviewer for these observations.

TABLE 6

Results of Regressions of Pooled *ALLP* on Abnormal Fee Measures

Variable	Abnormal Total Fee	Abnormal Nonaudit Fee
Intercept	−0.0005	0.0004
	(−0.87)	(0.61)
EXEMPT	−0.0001	−0.0001
	(−1.35)	(−1.37)
UTOTFEE	0.0000	
	(0.34)	
EXEMPT × *UTOTFEE*	−0.0005***	
	(−2.59)	
UNAFEE		−0.0001*
		(−1.69)
EXEMPT × *UNAFEE*		−0.0003**
		(−2.54)
BIG5	0.0000	0.0000
	(0.40)	(0.33)
MB	0.0000	0.0000
	(1.26)	(1.26)
LMVE	−0.0001*	−0.0001*
	(−1.90)	(−1.75)
LOSS	−0.0003	−0.0002
	(−0.17)	(−0.11)
PASTLLP	0.1345***	0.1327***
	(4.36)	(4.33)
EBP	0.0501***	0.0512***
	(2.98)	(3.00)
TIER1	0.0000	0.0000
	(0.19)	(0.21)
TCAP	0.0000	0.0000
	(0.04)	(0.10)
Year Controls	Yes	Yes
n	1618	1618
Adj. R^2	9.27%	10.03%

*, **, *** Indicate 0.10, 0.05, and 0.01 significance levels, respectively, for a two-tailed test.

The regression model is:

$$ALLP = \gamma_0 + \gamma_1 EXEMPT + \gamma_2 UFEE + \gamma_3 EXEMPT \times UFEE + \gamma_4 BIG5 + \gamma_5 MB + \gamma_6 LMVE + \gamma_7 LOSS + \gamma_8 PASTLLP + \gamma_9 EBP + \gamma_{10} TIER1_{t-1} + \gamma_{11} TCAP_{t-1} + YEARCONTROLS + \varepsilon.$$

t-statistics, in parentheses, are based on standard errors adjusted for firm-level clustering.

Please see Table 1 for variable definitions.

SOX became effective in 2004 for accelerated filers—i.e., firms with a public float of $75 million or more. We re-estimate model (3) for the years 2004 through 2006 by replacing *EXEMPT* with *NF* (non-accelerated filers), which equals 1 for banks that are exempt from both Section 404 of SOX and Section 36 of the FDI Act, and 0 otherwise. The results of this specification are in Table 7.

TABLE 7

Results of Regressions of Abnormal *LLP* on Abnormal Fee Measures for Pre- and Post-2004 Periods

Panel A: Income-Increasing (Negative) *ALLP*

	Years 2000–2003		Years 2004–2006	
Variable	**Abnormal Total Fee**	**Abnormal Nonaudit Fee**	**Abnormal Total Fee**	**Abnormal Nonaudit Fee**
(1)	**(2)**	**(3)**	**(4)**	**(5)**
Intercept	−0.0014**	−0.0015**	−0.0030**	−0.0029**
	(−2.02)	(−2.14)	(−2.46)	(−2.34)
SMALL/NF	−0.0002*	−0.0002	−0.0002	−0.0003
	(−1.92)	(−1.59)	(−1.05)	(−1.40)
UTOTFEE	−0.0001			
	(−0.66)			
SMALL × UTOTFEE	−0.0006**			
	(−1.99)			
UNAFEE		−0.0001		
		(−0.36)		
SMALL × UNAFEE		−0.0004*		
		(−1.84)		
UTOTFEE			−0.0000	
			(−0.06)	
NF × UTOTFEE			−0.0008**	
			(−2.50)	
UNAFEE				−0.0001
				(−0.90)
NF × UNAFEE				−0.0007**
				(−2.49)
BIG5	−0.0001	−0.0001	0.0002	0.0002
	(−1.58)	(−1.52)	(1.63)	(1.49)
MB	−0.0000	−0.0000	0.0002*	0.0002*
	(−1.34)	(−1.34)	(1.68)	(1.78)

(continued on next page)

Panel A: Income-Increasing (Negative) *ALLP*

Variable	Years 2000–2003		Years 2004–2006	
	Abnormal Total Fee	Abnormal Nonaudit Fee	Abnormal Total Fee	Abnormal Nonaudit Fee
(1)	(2)	(3)	(4)	(5)
LMVE	0.0001**	0.0001**	0.0000	0.0000
	(1.98)	(2.05)	(0.72)	(0.53)
LOSS	0.0000	0.0001	−0.0033***	−0.0033**
	(0.03)	(0.25)	(−2.69)	(−2.38)
PASTLLP	−0.0580**	−0.0592**	−0.0044	−0.0043
	(−2.09)	(−2.14)	(−0.17)	(−0.10)
EBP	−0.0147	−0.0138	−0.0233	0.0232
	(−1.47)	(−1.40)	(−1.33)	(−0.72)
TIER1	−0.0000	−0.0000	0.0001	0.0001
	(−0.23)	(−0.30)	(0.78)	(0.69)
TCAP	−0.0000	−0.0000	0.0000	0.0000
	(−0.31)	(−0.20)	(0.30)	(0.28)
Year Controls	Yes	Yes	Yes	Yes
n	457	457	425	425
Adj. R^2	3.47%	4.31%	11.03%	11.58%

Panel B: Income-Decreasing (Positive) *ALLP*

Variable	Years 2000–2003		Years 2004–2006	
	Abnormal Total Fee	Abnormal Nonaudit Fee	Abnormal Total Fee	Abnormal Nonaudit Fee
(1)	(2)	(3)	(4)	(5)
Intercept	0.0033**	0.0026**	0.0009	0.0007
	(2.36)	(2.67)	(1.07)	(0.78)
SMALL/NF	0.0001	0.0003*	0.0000	0.0000
	(0.84)	(1.69)	(0.27)	(0.20)

(continued on next page)

Panel B: Income-Decreasing (Positive) *ALLP*

	Years 2000–2003		Years 2004–2006	
Variable	**Abnormal Total Fee**	**Abnormal Nonaudit Fee**	**Abnormal Total Fee**	**Abnormal Nonaudit Fee**
(1)	(2)	(3)	(4)	(5)
UTOTFEE	0.0003**			
	(2.11)			
SMALL × UTOTFEE	−0.0009**			
	(−2.29)			
UNAFEE		−0.0000		
		(−0.04)		
SMALL × UNAFEE		−0.0003**		
		(−2.40)		
UTOTFEE			−0.0001	
			(−1.23)	
NF × UTOTFEE			−0.0001	
			(−0.38)	
UNAFEE				−0.0001*
				(−1.80)
NF × UNAFEE				−0.0002
				(−1.33)
BIG5	0.0001	0.0002	−0.0000	−0.0001
	(1.28)	(1.37)	(−0.19)	(−0.50)
MB	−0.0001**	−0.0001**	−0.0000	−0.0000
	(−2.13)	(−2.12)	(−0.44)	(−0.48)
LMVE	−0.0001**	−0.0001*	−0.0001**	−0.0001*
	(−2.77)	(−1.85)	(−1.99)	(−1.92)
LOSS	0.0019**	0.0018**	0.0005	0.0003
				(0.32)
	(2.36)	(2.14)	(0.59)	
PASTLLP	0.0555	0.0618*	0.1924***	0.1863**
	(1.53)	(1.66)	(2.90)	(2.35)

(continued on next page)

Panel B: Income-Decreasing (Positive) *ALLP*

	Years 2000–2003		Years 2004–2006	
Variable	**Abnormal Total Fee**	**Abnormal Nonaudit Fee**	**Abnormal Total Fee**	**Abnormal Nonaudit Fee**
(1)	(2)	(3)	(4)	(5)
EBP	0.0433**	0.0414**	0.0601***	0.0631***
	(2.33)	(2.03)	(3.32)	(3.30)
TIER1	0.0000	0.0000	0.0000	0.0000
	(0.30)	(0.68)	(0.28)	(0.16)
TCAP	−0.0001	−0.0001*	0.0000	0.0000
	(−1.44)	(−1.69)	(0.35)	(0.52)
Year Controls	Yes	Yes	Yes	Yes
n	349	349	387	387
Adj. R^2	12.72%	11.16%	14.98%	16.13%

*, **, *** Indicate 0.10, 0.05, and 0.01 significance levels, respectively, for a two-tailed test.
t-statistics, in parentheses, are based on standard errors adjusted for firm-level clustering.
The regression model is:

$$ALLP = \gamma_0 + \gamma_1 SMALL/NF + \gamma_2 UFEE + \gamma_3 SMALL/NF \times UFEE + \gamma_4 BIG5 + \gamma_5 MB + \gamma_6 LMVE + \gamma_7 LOSS + \gamma_8 PASTLLP + \gamma_9 EBP + \gamma_{10} TIER1_{t-1} + \gamma_{11} TCAP_{t-1} + YEARCONTROLS + \varepsilon.$$

Please see Table 1 for variable definitions.

We report results for the pre- and post-Section 404 of SOX periods. For the pre-Section 404 period (years 2000 through 2003), we interact *SMALL* with *UTOTFEE* and *UNAFEE*. For the post-Section 404 period (years 2004 through 2006) we interact *NF* with *UTOTFEE* and *UNAFEE*. Table 7, Panel A reports the results for income-increasing (negative) *ALLP* and Panel B reports the results for income-decreasing (positive) *ALLP*. We find that for income-increasing *ALLP*, the sum of the coefficients on *UTOTFEE* and *SMALL* × *UTOTFEE* and on *UNAFEE* and *SMALL* × *UNAFEE* are negative and significant at the 0.05 level. We also find that, for income-decreasing (positive) *ALLP*, the coefficients on *SMALL* × *UTOTFEE* and *SMALL* × *UNAFEE* are negative and significant at the 0.05 level. These results are consistent with the results in Tables 4 and 5, suggesting that there is greater earnings management via abnormal LLP by small banks (less than $500 million in total assets) that pay higher abnormal fees to the auditor. Turning to the post-Section 404 period, we find that, for income-increasing *ALLP*, the coefficients on *NF* × *UTOTFEE* and *NF* × *UNAFEE* are both negative and significant at the 0.05 level. We do not find a significant coefficient for income-decreasing *ALLP*. Overall, these results suggest that banks that are exempt from both Section 36 of the FDI Act and Section 404 of SOX and pay higher abnormal fees to the auditor engage in greater income-increasing earnings management via *ALLP*.

Sensitivity Tests

We conduct several sensitivity checks to assess the robustness of these results. First, we summarize the results for audit fees and fee ratio reported in Tables 4 and 5. Recall that our main results are based on unexpected total fees and unexpected nonaudit fees; however, consistent with prior research on auditor fees for industrial firms (Ashbaugh et al. 2003; Srinidhi and Gul 2007), we present results for audit fees and fee ratio as additional evidence. For income-increasing (negative) *ALLP*, neither abnormal audit fees nor actual audit fees is significantly associated with *ALLP*. On the other hand, unexpected fee ratio has a strong negative ($p < 0.01$) relation with income-increasing *ALLP* for small banks. For income-decreasing (positive) *ALLP*, actual audit fees have a positive ($p < 0.05$ or better) association with *ALLP* for both large and small banks; however, this relation does not hold for unexpected audit fees. Abnormal fee ratio is negatively associated with income-increasing *ALLP* for small banks ($p < 0.01$), but fee ratio is not significantly associated with income-increasing *ALLP*.

Second, we discuss the results (untabulated) for separate subsample analyses of small and large banks. For income-increasing *ALLP*, our main results for unexpected total fees and unexpected audit fees hold for the subsample analysis. More specifically, for large (non-exempted) banks, there is no significant association between unexpected total fees and unexpected nonaudit fees and income-increasing *ALLP*. For small (exempted) banks, the estimated coefficient for unexpected total fees is -0.0006 ($p < 0.05$), and for unexpected nonaudit fees the estimated coefficient is -0.0002 ($p < 0.05$), consistent with our main results reported in Table 4, Panel A. For income-decreasing *ALLP*, the only significant association is with unexpected nonaudit fees for small (exempted) banks (estimated coefficient of -0.0003, $p < 0.01$), consistent with the main results reported in Table 5, Panel A.

Third, we assess the sensitivity of our results to omitting observations that have an abnormal influence on our regressions. When we delete observations that have R-student values greater than 3 or less than -3, the results for unexpected total fees and unexpected nonaudit fees for both income-increasing and income-decreasing *ALLP* are similar to the results reported in Panel A of Tables 4 and 5.[32]

[32] Instead of deleting the top 1 percent of observations, when we winsorize the key variables at the top 1 percent for income-increasing *ALLP*, the coefficient on *EXEMPT* × *UTOTFEE* is -0.0007 (significant at the 0.01 level). Similarly,

Fourth, we examine the robustness of our results to interacting the capital management and income smoothing variables with *EXEMPT*. Using income-increasing *ALLP* observations, we re-estimate model (3) with the following interactions: *EXEMPT* × *EBP*, *EXEMPT* × *TIER1*, and *EXEMPT* × *TCAP*. Untabulated results indicate that neither the capital ratios nor any of their interactions are significant. Additionally, the income smoothing variable and its interaction are also not significant. However, the coefficients on *EXEMPT* × *UTOTFEE* and *EXEMPT* × *UNAFEE* continue to be negative and significant at the 0.05 level.

Finally, we examine the relation between abnormal LLP and auditor fees using an alternate estimate of abnormal LLP. Ahmed et al. (1999) suggest that the earnings management through LLP reported in prior research is conditional on the inclusion of beginning non-performing loans (NPL_{t-1}) in the abnormal LLP model. Our main results are robust to using estimates of abnormal LLP after excluding NPL_{t-1} from model (1).

VI. CONCLUSIONS

Despite the economic importance of the banking industry, there is limited research on the various relationships that exist between banks and their auditors. We provide empirical evidence on the relation between fees paid to the auditor and the extent of earnings management via loan loss provisions in the banking industry. In particular, we study a setting where small banks are exempt from regulatory scrutiny. Section 36 of the FDI Act, which was intended to aid early detection of problems in the financial management of insured banks, requires an assessment of the effectiveness of internal control over financial reporting by auditors. However, this requirement does not apply to banks with total assets of less than $1 billion (less than $500 million until 2004). We study whether there is a relation between earnings management via loan loss provisions and fees paid to the auditor, particularly for banks that are exempt from the FDI Act.

We report several important findings. First, for large banks, we find that both unexpected total fees and unexpected nonaudit fees are not related to either income-increasing or income-decreasing *ALLP*. These findings indicate that unexpected auditor fees are unrelated to earnings management behavior in large banks, suggesting that auditor fee dependence is not a threat to auditor independence in large banks. In contrast, we find a strong negative association between income-increasing (negative) *ALLP* and both unexpected total fees and unexpected nonaudit fees for small banks that are exempt from Section 36 of the FDI Act and Section 404 of SOX. Furthermore, we also document a strong negative association between unexpected nonaudit fees and income-decreasing (positive) *ALLP*. These results are consistent with greater earnings management by exempt banks that pay higher abnormal fees to the auditor. Collectively, our results suggest that auditor independence could be threatened when auditor fee dependence on the audit client is higher, but only for small banks that are not subject to the same level of regulatory scrutiny as larger banks.

Our results complement the findings of Altamuro and Beatty (2010), who link FDICIA requirements to improved loan loss provisioning. We also find that loan loss provisions of banks that are subject to internal control requirements under FDICIA or SOX are independent of fees paid to auditors. Additionally, we document that banks that are exempted from internal control requirements under FDICIA or SOX are more likely to overstate earnings when the fees paid to auditors are high.

Our results have several implications. First, they inform policymakers that auditor fees are related to earnings management for small but not large banks. This suggests that any potential

---

the coefficient on *EXEMPT* × *UNAFEE* is −0.0003 (significant at the 0.01 level). For income-decreasing *ALLP*, the coefficient on *EXEMPT* × *UNAFEE* is −0.0002 (significant at the 0.05 level). Once again, we find that unexpected total and nonaudit fees are unrelated to *ALLP* for non-exempted banks.

impairment of auditor independence through LLP could be a problem for small banks. Second, to the extent that the FDICIA can be viewed as a precursor to SOX and to the extent that our results for banks are generalizable to other industries, our findings of increased earnings management by small banks that are less closely regulated are relevant to the debate on reducing the requirements of SOX for small firms. Finally, our findings are timely and relevant given the recent banking crisis and the response from governments around the world contemplating new regulations. In particular, our results inform policymakers on the relationship that existed between fees paid to auditors and the extent of earnings management in banks prior to the current banking crisis.

REFERENCES

Ahmed, A. S., C. Takeda, and S. Thomas. 1999. Bank loan loss provisions: A reexamination of capital management, earnings management and signaling effects. *Journal of Accounting and Economics* 28 (November): 1–25.

Aiken, L., and S. West. 1991. *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage Publications Inc.

Altamuro, J., and A. Beatty. 2010. How does internal control regulation affect financial reporting? *Journal of Accounting and Economics* 49 (February): 58–74.

Ashbaugh, H., R. LaFond, and B. Mayhew. 2003. Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review* 78 (July): 611–639.

Beaver, W. H., and E. Engel. 1996. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics* 22 (August-December): 177–206.

Benston, G. J., and G. G. Kaufman. 1997. FDICIA after five years. *The Journal of Economic Perspectives* 11 (Summer): 139–158.

Boritz, J. E., and P. Zhang. 1997. The implications of alternative litigation cost allocation systems for the value of audits. *Journal of Accounting, Auditing & Finance* 12 (Fall): 353–372.

Collins, J. H., D. A. Shackelford, and J. M. Wahlen. 1995. Bank differences in the coordination of regulatory capital, earnings, and taxes. *Journal of Accounting Research* 33 (Autumn): 263–291.

De, S., and P. K. Sen. 2002. Legal liabilities, audit accuracy and the market for audit services. *Journal of Business Finance & Accounting* 29 (April/May): 353–410.

Edwards, F. R., and F. S. Mishkin. 1995. The decline of traditional banking: Implications for financial stability and regulatory policy. *Federal Reserve Bank of New York Economic Policy Review* 1 (2): 27–45.

Federal Deposit Insurance Corporation (FDIC). 1993. *Part 363-Annual Independent Audits and Reporting Requirements*. Available at: http://www.ffiec.gov/ffiecinfobase/resources/audit/fdi-12cfr363-annual_independ_audit_report_requir.pdf.

———. 2001. *Financial Institution Letters*. Available at: http://www.fdic.gov/news/news/financial/2001/fil0163.html.

———. 2005a. *Proposed Rules*. Available at: http://www.gpo.gov/fdsys/pkg/FR-2005-08-02/html/05-15109.htm.

———. 2005b. *Financial Institution Letters*. Available at: http://www.fdic.gov/news/news/financial/2005/fil11905a.html.

Fields, L. P., D. R. Fraser, and M. S. Wilkins. 2004. An investigation of the pricing of audit services for financial institutions. *Journal of Accounting and Public Policy* 23 (January/February): 53–77.

Firth, M. 1997. Provision of nonaudit services by accounting firms to their clients. *Contemporary Accounting Research* 14 (Summer): 1–21.

Gao, F., J. S. Wu, and J. Zimmerman. 2009. Unintended consequences of granting small firms exemptions from securities regulation: Evidence from the Sarbanes-Oxley Act. *Journal of Accounting Research* 47 (May): 459–506.

Guay, W. R., S. P. Kothari, and R. L. Watts. 1996. A market-based evaluation of discretionary accrual models. *Journal of Accounting Research* 34 (Spring): 83–105.

Jones, K. L., G. V. Krishnan, and K. Melendrez. 2008. Do discretionary accruals models detect actual cases

of fraudulent and restated earnings? An empirical evaluation. *Contemporary Accounting Research* 25 (Summer): 499–531.

Kanagaretnam, K., G. Lobo, and R. Mathieu. 2003. Managerial incentives for income smoothing through loan loss provisions. *Review of Quantitative Finance and Accounting* 20: 63–80.

———, ———, and D. Yang. 2004. Joint tests of signaling and income smoothing through bank loan loss provisions. *Contemporary Accounting Research* 21 (Winter): 843–884.

Kim, J. B., X. Liu, and D. A. Simunic. 2008. Audit pricing, legal liability regimes, and Big 4 premiums: Theory and cross-country evidence. *Contemporary Accounting Research* 25 (Spring): 55–99.

Kinney, W. R., Jr., and R. Libby. 2002. Discussion of the relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77 (Supplement): 107–114.

Libby, R. 1981. *Accounting and Human Information Processing: Theory and Applications*. Englewood Cliffs, NJ: Prentice Hall, Inc.

McNichols, M. F. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19 (4–5): 313–345.

———. 2002. Discussion of: The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 61–69.

Murphy, C. W. 2004. *Goodbye FDICIA, Hello Sarbox—Sort Of*. New York, NY: PricewaterhouseCoopers Publications.

Neter, J., W. Wasserman, and M. H. Kutner. 1989. *Applied Linear Regression Models*. 2nd edition. Homewood, IL: Irwin.

Phillips, L. J., and D. G. Lierley. 2001/2002. Allowance for loan losses guidance: Where we stand today. *Commercial Lending Review* 17 (Winter): 7–10.

Runkel, P., and J. McGrath. 1972. *Research on Human Behavior—A Systematic Guide to Method*. New York, NY: Holt, Rinehart, and Winston, Inc.

Srinidhi, B. N., and F. A. Gul. 2007. The differential effects of auditors' nonaudit and audit fees on accrual quality. *Contemporary Accounting Research* 24 (Summer): 595–629.

St. Pierre, K., and J. A. Anderson. 1984. An analysis of the factors associated with lawsuits against public accountants. *The Accounting Review* 59 (April): 242–263.

Wahlen, J. M. 1994. The nature of information in commercial bank loan loss disclosures. *The Accounting Review* 69 (July): 455–478.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 2047–2074

American Accounting Association
DOI: 10.2308/accr.2010.85.6.2047

Do Supplementary Sales Forecasts Increase the Credibility of Financial Analysts' Earnings Forecasts?

Edmund C. Keung
National University of Singapore

ABSTRACT: This study examines whether the market reacts more strongly to earnings forecast revisions when financial analysts supplement their earnings forecasts with sales forecasts. I find that earnings forecast revisions supplemented with sales forecast revisions have a greater impact on security prices than do stand-alone earnings forecast revisions, controlling for the incremental information content in sales forecasts. Supplemented earnings forecasts are more accurate *ex post*, controlling for other individual analyst characteristics. Results are robust to controlling for earnings persistence and time effects. Taken as a whole, financial analysts are more likely to supplement their earnings forecasts with sales forecasts when they have better information. Supplementary sales forecasts appear to lend credibility to earnings forecasts because financial analysts provide sales forecasts when they are more informed.

Keywords: *financial analyst; sales forecast; revenue forecast; forecast accuracy.*

Data Availability: *All data are available from the sources identified in the text.*

I. INTRODUCTION

This study investigates two research questions: (1) How does the market react to earnings forecast revisions when individual financial analysts justify their earnings forecasts with sales forecasts? (2) Do financial analysts provide supplementary sales forecasts when they are better informed? I examine whether investors react more strongly to earnings forecast revisions supplemented with sales forecast revisions (hereafter "supplemented earnings forecast revisions") than to revisions not supplemented with sales forecast revisions (hereafter "stand-alone earnings forecast revisions"). I define an earnings forecast revision as *supplemented* by a sales

This paper is based on part of my dissertation at Washington University. I am indebted to Rich Frankel, chairman of my dissertation committee, for his guidance. I thank Steven Kachelmeier (editor), Mark DeFond, Dan Dhaliwal, Peter Easton, Paul Healy, Hung Ming-Yi, Wayne Landsman, Steve Matsunaga, Mujtaba Mian, Mikhail Pevzner, Srini Sankaraguruswamy, Sheng Jianfeng, Michael Shih, Grantley Taylor, Trevor Wilkins, Oscar Yee, Eric Yeung, Tzachi Zach, Ian Zimmer, and two anonymous referees for their comments and suggestions. I have also benefited from comments on an earlier version presented at the 2007 AAA Annual Meeting and the 2007 AFAANZ Conference. All errors are mine. I acknowledge the contribution of Thomson Financial for providing earnings per share forecast and sales forecast data, available through the Institutional Brokers Estimate System (I/B/E/S).

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: February 2008
Accepted: April 2010
Published Online: November 2010

forecast if a financial analyst revises both his/her earnings and sales forecasts on the same day as reported in the Institutional Brokers Estimate System (I/B/E/S) database[1]

I define forecast credibility as the extent to which investors believe and react to earnings forecasts. I measure forecast credibility using the forecast response coefficient to an earnings forecast revision. Supplementary sales forecasts appear to increase the credibility of earnings forecasts for two reasons. First, investors can better evaluate the credibility of earnings forecasts with their knowledge of the business and other macroeconomic factors when financial analysts reveal sales forecasts. Sales are an important driver of firm value and a key variable to investors engaged in fundamental analysis. For example, Penman (2001, 503) suggests that sales forecasts contain information on sales growth and the firm's marketing plan, which is not contained in earnings forecasts. Financial analysts convey the causes of the earnings forecast revisions through sales forecasts. Sales forecasts explain whether the revisions in earnings forecast are due to growth in revenue, increase in profit margin, or cost control. Earnings growth supported by an increase in revenue is more sustainable because growth in revenue reflects the success of the underlying product differentiation strategy (Porter 1980). Second, sales forecasts are *ex post* verifiable at the year-end.[2] Because investors can use sales forecasts to gauge the information provided by individual financial analysts, financial analysts are more likely to issue supplementary sales forecasts when they are more informed.[3]

While prior research mainly focuses on financial analysts' earnings forecasts and their stock recommendations, it is an open question as to how the market reacts to other supplementary information provided by financial analysts.[4] My first set of tests examines the association between supplementary sales forecast revisions and the market response to financial analysts' earnings forecast revisions.[5] On the one hand, if supplementary sales forecasts add credibility to earnings forecast revisions, then price reactions to the earnings forecast revisions would be stronger, controlling for the information content of sales forecasts. On the other hand, if the market perceives supplementary sales forecasts as "cheap talk" to generate trading volume, then sales forecasts and price reactions would not be associated. My results show that supplemented earnings forecast revisions have a greater impact on contemporaneous stock prices than do stand-alone earnings forecast revisions, controlling for the independent information in sales forecast revisions and other individual analyst characteristics. The difference in market responses suggests that investors view supplemented earnings forecast revisions as more credible.

I then examine whether the financial analysts have better information when they choose to reveal their supplementary sales forecasts. My second set of tests examines the accuracy of supplemented earnings forecasts, controlling for other individual analyst characteristics. Because analysts' private information is unobservable, I control for other confounding factors in two ways. First, I measure forecast accuracy relative to other analysts issuing forecasts for the same firm-year to control for potential differences in forecasting difficulty across firm-years (Clement and Tse

---

[1] For the purposes of this study, sales forecast and revenue forecast carry the same meaning and are used interchangeably unless otherwise stated.

[2] *Ex post* verifiable forecasts (*hard* forecasts) are those forecasts that the actual values are observed by the market after a definite period of time, for example, at subsequent earnings announcement dates. In contrast to *hard* forecasts, *soft* forecasts are those that are not specific enough to be compared with subsequent realizations (Hutton et al. 2003). An example of a soft forecast would be a qualitative discussion of a firm's prospects in an analyst report.

[3] For example, Loh and Mian (2006) argue that quality of justification of earnings forecasts in analyst reports is part of an analyst's utility function.

[4] Call et al. (2009) is an exception. However, they do not examine the information content of and the market reactions to cash flow forecasts.

[5] Because forecasting sales is the first step to forecasting earnings, it is unlikely that an analyst would maintain the same sales forecast if he/she revises his/her own earnings forecast. Therefore, I assume that it is a decision by individual financial analysts whether to disclose their revised sales forecasts.

2005; Herrmann and Thomas 2005). Second, I directly control for other individual analyst and forecast characteristics (Mikhail et al. 1997; Clement 1999; Jacob et al. 1999; Brown 2001; Clement and Tse 2003). Consistent with my hypotheses, I find that supplementary sales forecasts are associated with more accurate earnings forecasts. Financial analysts are more likely to supplement their earnings forecasts with sales forecasts when they have better information. These results also suggest that supplementary sales forecasts lend credibility to earnings forecasts because financial analysts provide sales forecasts when they are more informed.

This study extends the findings of Baginski et al. (2004) who demonstrate that management earnings forecasts supplemented by sales forecasts receive larger price responses, which is consistent with investors perceiving these earnings forecasts as more credible. Because managers and analysts have different incentives and differential access to information, I investigate whether Baginski et al.'s (2004) results extend to analysts' earnings forecasts.

It is not self-evident that investors will have identical interpretations of managers' and analysts' forecasts. First, firm mangers and financial analysts belong to different labor markets and are subject to different incentives. On the one hand, management compensation and career outcome are directly tied to the performance of the company. While supplementary information could be useful to the market, the primary concern of managers is not to convey their forecast accuracy. If the company is not performing well, then management would have incentives to avoid shareholder discipline by attributing the poor performance to factors that are beyond their control. On the other hand, firm performance does not have a direct effect on the payoff of financial analysts. They are compensated for the amount of trading commissions generated. As a result, the market may *not* react to supplementary information if the market perceives such supplementary sales forecast to be merely cheap talk.

Second, firm mangers and financial analysts have differential access to information. While financial analysts have a competitive edge in understanding industry information, firm managers have access to firm-specific proprietary information related to the firms they are running. In other words, they have first-hand information relative to the financial analysts and have discretion on how much information (if any) is to be released to financial analysts. *A priori*, it is not obvious whether supplementary information provided by firm mangers and financial analysts is equally reliable.[6]

Baginski et al. (2004) regress three-day cumulative abnormal returns on the interaction between unexpected earnings and an indicator variable for the presence of sales forecasts. However, they do not control for the incremental information content in the sales forecasts. I include *both* revisions in sales forecasts and indicators for the presence of sales forecasts to disentangle whether the market reactions are driven by: (1) specific information in sales forecasts, and/or (2) the act of disclosing sales forecasts. My study also goes one step further and investigates whether these supplemented earnings forecasts are indeed more accurate *ex post*. Because the role of financial analysts has increased dramatically over time, it is important to understand the nature of and the market responses to supplementary sales forecasts.

Givoly et al. (2009) provide evidence that cash flow forecasts appear to be a naive extension of analysts' earnings forecasts. They show that analysts' forecasts of cash flows are of limited information content and are only weakly associated with stock returns. My evidence suggests that sales forecasts, unlike cash flow forecasts, provide incremental information to the market.[7]

---

[6] I thank Steve Matsunaga for sharing his insights on the differences between management forecasts and analysts' forecasts with me.

[7] One possible reason is that from a financial statement analysis perspective, forecasting sales is the first step to forecasting earnings (for example, see Penman [2001], among others). In other words, financial analysts derive their earnings forecasts *from* their sales forecasts. Givoly et al.'s (2009) evidence suggests cash flow forecasts appear to be a naive

My work is related to Rees and Sivaramakrishnan (2007).[8] However, this study differs in two ways. First, while Rees and Sivaramakrishnan (2007) examine the implication of meeting or exceeding revenue forecasts at the time of the quarterly earnings announcement, I do not restrict my sample to the last forecast revisions issued by financial analysts and look at revisions of both earnings and supplementary sales forecasts *throughout* the year. Because individual analyst forecasts are a vital source of timely information for investors, it is important to examine how supplementary information affects the immediate price impact of forecast revisions. Second, I examine *why* the market reacts more strongly to supplemented earnings forecasts. I find that supplemented earnings forecasts are more informative because supplementary sales forecasts are a credible signal of analysts' accuracy.

My study contributes to the literature in three ways. First, I extend the literature on the role of and the market reactions to information that supplements earnings forecasts. I examine supplementary sales forecasts, a relatively new form of forecast disseminated by analyst tracking services. Second, I rule out the competing hypothesis that sales forecasts are uninformative cheap talk provided by financial analysts, merely hoping to stimulate trading volume and generate trading commissions. Third, I provide empirical evidence that the presence of supplementary sales forecasts, conditional on other individual analyst and forecast characteristics, could be used to identify the more accurate and informative earnings forecasts.

Section II discusses prior literature and develops hypotheses. Section III describes sample selection criteria and provides sample descriptive statistics. Sections IV and V present the empirical results for tests of market reaction and tests of accuracy of supplemented versus stand-alone earnings forecasts. Section VI presents the results for additional robustness tests. Section VII concludes.

II. RELATED LITERATURE AND HYPOTHESES DEVELOPMENT

Prior research documents the market responses to individual analyst earnings forecast revisions (Lys and Sohn 1990; Stickel 1991; Park and Stice 2000).[9] While individual analyst earnings forecasts are informative, earnings forecasts are only one of the many signals about firms' prospects provided by financial analysts. Francis and Soffer (1997) find that earnings forecasts and stock recommendations are incrementally informative to one another. When stock recommendations are favorable, investors place greater reliance on earnings forecast revisions, which are *ex post* verifiable. Bradshaw (2002) examines a sample of analyst reports and documents the frequency at which analysts disclose target prices as justifications for their stock recommendations. He finds that the most favorable recommendations and target prices are more likely to be justified by quantitative forecasts such as price-earnings ratios and expected growth, while the least favorable recommendations are more likely to be justified with other qualitative statements. These findings suggest that, when favorable stock recommendations are perceived as less credible, financial analysts have the incentives to justify their recommendations.

Supplementary sales forecasts could be useful in two ways. First, they convey information about the prospects of the firms, independent of the information in earnings forecasts. Second,

---

extension of analysts' earnings forecasts, i.e., financial analysts derive their cash flow forecasts from their earnings forecasts.

[8] Rees and Sivaramakrishnan (2007) examine whether meeting or exceeding revenue forecasts serves as an additional cue to the market in pricing earnings performance. They show that meeting revenue forecasts is of greater importance to market participants than the magnitude of the revenue forecast error *per se*.

[9] Lys and Sohn (1990) find that individual earnings forecasts are informative to the market. Stickel (1991) find that earnings forecast revisions affect stock prices, but prices do not immediately assimilate the information in the revisions. Prices continue to drift in the direction of the revision for about six months after the revision. Park and Stice (2000) find that current revision response coefficient is increasing in the analyst's past earnings forecast accuracy.

supplementary sales forecasts increase the credibility of earnings forecasts. Asquith et al. (2005) examine the market reactions to analyst reports, finding that market reactions are associated with both the magnitude of earnings forecast revision and the strength of analysts' arguments in the analyst reports.[10] Although Asquith et al. (2005) show that supplementary information in analyst reports provides incremental information to the market, they do not distinguish whether the supplementary information provided by financial analysts conveys information independent of earnings forecasts or whether it adds credibility to earnings forecasts.

Hutton et al. (2003) find that good news management earnings forecasts are informative only when supplemented with verifiable forward-looking statements. Baginski et al. (2004) investigate market reactions to explanation (attribution) for the management forecasts. They classify attributions into two types: external and internal. Internal attributions are causal factors internal to the organization, such as strategy or management skill. External attributions are events outside the company, such as competition and certain economic factors. They find a positive relation between security price reaction and management forecast with external attribution, but not for internal attribution. Wasley and Wu (2006) find that managers issue cash flow forecasts to signal good news in cash flows and to lend credibility to good news in earnings. Taken together, these findings suggest that verifiable supplementary information lends credibility to earnings forecasts. Because supplementary sales forecasts justify earnings forecasts and are *ex post* verifiable, I hypothesize that they increase the credibility of earnings forecasts. Thus, my first hypothesis is as follows:

H1: Supplementary sales forecasts add credibility to earnings forecast revisions.

Perceived forecast accuracy increases with analyst reputation (Stickel 1992; Sinha et al. 1997).[11] Clement and Tse (2003) find the market reaction is stronger when individual analyst characteristics predict that the earnings forecast revisions are more likely to be accurate *ex ante*. Prior literature finds that individual analyst earnings forecast accuracy increases with prior accuracy, brokerage size, forecast frequency, and firm experience, and decreases with days elapsed since the last forecast, forecast horizon, number of companies, and number of industries the financial analyst follows (Mikhail et al. 1997; Clement 1999; Jacob et al. 1999; Brown 2001).[12] The market learns about the ability of financial analysts and updates its beliefs about their perceived accuracy through their performance and action over time (Chen et al. 2005).

As more information about analyst ability is revealed, the market reacts more strongly, resulting in higher trading volume (Abarbanell et al. 1995). Hayes (1998) shows that trading volume

---

[10] Asquith et al. (2005) measure the level of analyst justification by aggregating the number of positive remarks less the number of negative remarks related to 14 specific criteria about the firm in the analyst report: revenue growth, earnings growth, new product introductions, new projects, cost efficiencies, expectations met, mergers and acquisitions, repurchase programs, industry climate, management, international operations, leverage, competition, and risk. I go through a sample of analyst reports that include revenue forecasts and those that do not. I find both types of analyst reports are not substantially different from each other in terms of structure and length. While most reports discuss sales growth and other quantitative or qualitative analysis supporting the EPS forecasts and stock recommendation, financial analysts often do not explicitly disclose their revenue forecasts in their reports. Some reports contain detailed sales forecasts by geographic segments and product lines, and explain how these sales forecasts translate into the EPS forecasts, assuming different profit margins across different segments. Operating cash flow forecasts are not as common as sales forecasts and are often provided only when the financial analysts estimate the target price using some cash flow-based models.

[11] Stickel (1992) defines that an analyst has a superior reputation if s/he is included in the Institutional Investor All-American Research Team. He shows that earnings forecast revisions by All-American analysts have a larger impact on security prices than do revisions by other analysts. Sinha et al. (1997) use past forecast accuracy to identify superior analysts and find that superior forecasting ability persists over time.

[12] Mikhail et al. (1997) document a decline in the absolute value of analyst quarterly forecast errors as firm-specific experience increases. Clement (1999) finds that forecast accuracy is positively associated with analysts' experience and employer size, but is negatively associated with the numbers of firms and industries followed by the analyst. Jacob et al. (1999) report that analysts' aptitude and brokerage house characteristics are associated with forecast accuracy. Brown (2001) finds that past performance predicts forecast accuracy. Clement and Tse (2003) find that forecast frequency is as important as past accuracy in explaining future accuracy.

increases in the perceived accuracy of the forecast. Because financial analysts are rewarded for the amount of trading commissions they generate for their brokerage houses, I argue that the more able analysts have incentives to convey their ability so that the market will trade on their analyst reports (Schipper 1991; Jackson 2005).[13] On the one hand, if financial analysts provide sales forecasts to enhance credibility when they are more informed, then the presence of supplementary sales forecasts should be associated with more accurate earnings forecasts, controlling for other determinants of forecast accuracy. On the other hand, if supplementary sales forecasts are issued by uninformed analysts hoping to stimulate trading volume and generate trading commissions, then supplementary sales forecasts would have no association with earnings forecast accuracy *ex post*. I examine whether supplementary sales forecasts, which are provided to the investors at the discretion of financial analysts, could be an *ex ante* signal to the market that the financial analysts are better informed. Thus, my second hypothesis is as follows:

H2: Earnings forecasts supplemented with sales forecast revisions are more accurate than stand-alone earnings forecasts.

III. SAMPLE SELECTION AND DESCRIPTIVE STATISTICS

Sample Selection

I obtain data on analyst forecasts and individual analyst characteristics from the Institutional Brokers Estimate System (I/B/E/S) database, financial statements data and earnings announcement dates from Compustat, and stock returns data from the Center for Research on Security Prices (CRSP) database. I restrict my analyses to annual forecasts to avoid the complication of seasonality. Forecast revisions in the initial sample are required to have corresponding daily security return data available in the CRSP database. I define an earnings (sales) forecast revision as the newly revised earnings (sales) forecast less the analyst's own prior earnings (sales) forecast.[14] A sales forecast revision is defined as supplementary when a financial analyst revises both his/her earnings and sales forecasts on the same day as reported in the I/B/E/S detail file. I measure the information content of a supplementary sales forecast using three-day size-adjusted buy-and-hold return around the revision date. Size-adjusted buy-and-hold return is the raw buy-and-hold return of the firm minus the mean buy-and-hold return of an equal-weighted portfolio of firms in the same NYSE/AMEX/NASDAQ size decile over the three-day holding period.

To allow for variation in individual analyst and revision characteristics at the time of the forecast revisions, I perform the empirical analysis at the forecast revision level. I remove earnings forecast revisions around quarterly earnings announcements (-1 to $+1$) because the price responses to the news in the earnings announcements might confound measures of price responses to earnings forecast revisions. I eliminate forecast revisions after the fiscal year-end to make my results consistent and comparable to other results documented in the literature (for example, Clement and Tse [2005], among others). These forecasts are likely to be confounded by earnings pre-announcements made by management (Soffer et al. 2000).[15] Following Gleason and Lee (2003), in order to mitigate inflated test statistics due to sample dependence induced by clustered

[13] This is known as the "soft dollar" practice, where institutional investors intentionally direct a portion of their trades through the brokerage houses whose financial analysts have provided the trade-generating research. Because brokerage firms' profits depend directly on commission revenues, analyst compensation (usually in the form of year-end bonuses) is implicitly tied to the amount of commissions they have generated from their research (Irvine 2000, 2004).

[14] Prior studies (for example, Stickel 1990; Park and Stice 2000; Gleason and Lee 2003) show that an analyst's own prior forecast is a better benchmark than the consensus forecast for measuring the amount of surprise in an individual analyst forecast revision. Results are robust to the choice of benchmark.

[15] Forecast revisions after the fiscal year account for 6 percent of all the individual annual earnings forecasts. My results are not affected by whether I include those forecasts.

revisions, I create a non-overlapping sample by including revisions that have no other revisions occurring within a three-day window. In other words, when more than one revision occurs in a three-day event window, I randomly pick only one of those forecast revisions. The non-overlapping sample contains 239,339 earnings forecast revisions from 1998 to 2003.[16] For regressions with individual analyst characteristics as independent variables, I randomly select one revision for an analyst if there are multiple revisions by the same analyst within a firm-year. This mitigates the concern that the sample is over-represented by multiple revisions from the same analyst. This reduces the sample to 132,146 earnings forecast revisions.[17]

Descriptive Statistics

Table 1 reports descriptive statistics on supplemented and stand-alone earnings forecast revisions. Both earnings and sales forecast revisions are deflated by stock prices two days before the revisions. On average, financial analysts revise their earnings forecasts downward (−0.0054 and −0.0057). This direction is consistent with firm managers' incentives to guide analysts' forecasts downward (Matsumoto 2002), so that firm managers could meet or beat consensus analyst forecasts. The average magnitude of earnings forecast revisions (0.0076) is smaller for supplemented earnings forecast revision than for stand-alone earnings forecast revisions (0.0086). However, supplemented earnings forecast revisions result in a stronger magnitude of market reactions (6.68 percent) than do stand-alone revisions (5.21 percent). Upward earnings forecast revisions are more common for supplemented revisions (43.6 percent) than for stand-alone revisions (40.6 percent). When an earnings forecasts revision is supplemented with a sales forecast revision, the market is more likely to react in the same direction as the earnings forecast revision. For example, 58.5 percent of supplemented earnings forecast revisions are associated with abnormal returns in the same direction as the earnings forecast revisions, as compared to 55.7 percent for stand-alone earnings forecast revisions.

Taken as a whole, the univariate results show that the market reacts more strongly and in a more consistent manner to the direction of earnings forecast revisions when they are supplemented with sales forecast revisions. However, these results do not distinguish whether supplementary sales forecasts convey information independent of earnings forecast revisions or add credibility to earnings forecasts. I examine this issue using multivariate regression analyses in the next section.

IV. MARKET REACTIONS TO SUPPLEMENTARY SALES FORECASTS

Research Design

In this section I test whether supplementary sales forecast revisions add credibility to earnings forecast revisions (H1) by examining whether investors react more strongly to them. I estimate the following regression (subscripts omitted):

$$BHAR3 = \beta_0 + \beta_1 E\_REV + \beta_2 E\_REV * SUPPL + \beta_3 S\_REV + \varepsilon \quad (1)$$

where:

[16] I start with the year 1998 because it is the first year with a significant number of annual sales forecasts. Pooled ordinary least-squares (OLS) regression is applied to the non-overlapping sample in the short-window price reactions analysis because by construction clustering of the observations is mitigated.

[17] This design choice is made to be conservative such that the statistical significance would not be overestimated. My results are robust to using the original overlapping sample or using the last forecast revisions issued by the financial analysts in the firm-years.

TABLE 1

Descriptive Statistics on Supplemented and Stand-Alone Earnings Forecast Revisions

	Supplemented Earnings Forecast Revisions			Stand-Alone Earnings Forecast Revisions			Two-Sample t-test	Wilcoxon Rank-Sum Test
	Mean	Std. Dev.	Median	Mean	Std. Dev.	Median	t-stat (p-value)	z-score (p-value)
Earnings forecast revision (*E_REV*)	−0.0054	0.0205	−0.0006	−0.0057	0.0205	−0.0008	−2.77 (<0.001)	−7.50 (<0.001)
Magnitude of earnings forecast revision (|*E_REV*|)	0.0076	0.0135	0.0026	0.0086	0.0028	0.0153	9.31 (<0.001)	−7.65 (<0.001)
Sales forecast revision (*S_REV*)	−0.0174	0.4182	−0.0002					
Magnitude of sales forecast revision (|*S_REV*|)	0.0290	0.0522	0.0096					
Three-day size-adjusted buy-and-hold return (*BHAR3*)	−0.43%	7.28%	−0.29%	−0.21%	6.14%	−0.20%	6.21 (<0.001)	4.18 (<0.001)
Magnitude of three-day size-adjusted buy-and-hold return (|*BHAR3*|)	6.68%	6.21%	4.69%	5.21%	5.34%	3.46%	35.49 (<0.001)	33.65 (<0.001)
Percentage of upward earnings forecast revisions	43.64%			40.64%			−10.89 (<0.001)	
Percentage of upward sales forecast revisions	58.76%							
Percentage of positive abnormal returns	47.64%			48.10%			1.63 (0.104)	
Percentage of observations where abnormal returns and earnings forecasts revision are in the same direction	58.46%			55.71%			−9.90 (<0.001)	
Number of earnings forecast revisions	28,004			211,335				

This table presents descriptive statistics on supplemented and stand-alone earnings forecast revisions. Revision in earnings (sales) forecast is the difference between the revised earnings (sales) forecast and the analyst's own prior earnings (sales) forecast deflated by stock price two days before revision. Size-adjusted return is the compounded buy-and-hold

(continued on next page)

TABLE 1 (continued)

return of the firm minus the compounded return for an equal-weighted portfolio of firms in the same NYSE/AMEX/NASDAQ size decile. I define a sales forecast revision as supplementary if a financial analyst revises both his/her earnings and sales forecasts on the same day as reported in the I/B/E/S detail file. Corresponding t-statistics for parametric two-sample t-tests and *z*-scores for nonparametric Wilcoxon rank-sum tests are reported.

BHAR3 = three-day size-adjusted buy-and-hold return (−1 to +1) surrounding an annual earnings forecast revision. Size-adjusted buy-and-hold return is the raw buy-and-hold return of the firm minus the mean buy-and-hold return of an equal-weighted portfolio of firms in the same NYSE/AMEX/NASDAQ size decile over the holding period;
E_REV = revision in annual earnings forecast relative to the analyst's own prior forecast deflated by stock price two days before revision;
SUPPL = 1 if the revision in earnings forecast is supplemented with revision in sales forecast, and 0 otherwise; and
S_REV = revision in annual sales forecast relative to the analyst's own prior forecast deflated by stock price two days before revision if it is revealed, and 0 otherwise.

In Equation (1), I examine whether the supplementary information provided by financial analysts conveys information independent of earnings forecast revisions, or whether such information conveys information on the credibility of the earnings forecasts. I include revision in sales forecast (*S_REV*) as a separate independent variable to control for the information in supplementary sales forecast revisions independent of earnings forecast revisions.[18] I use the interaction between earnings forecast revision (*E_REV*) and the indicator variable for the *presence* of supplementary sales forecast (*SUPPL*) to capture the incremental effect of supplementary sales forecast on the forecast response coefficient.

My methodology extends that used by Baginski et al. (2004). They regress three-day *CAR* on the interaction between unexpected earnings and an indicator variable for the presence of sales forecasts. However, they do not control for the incremental information content in the sales forecasts. Including *both* revisions in sales forecasts and indicators for the presence of sales forecasts allow me to disentangle whether the market reactions are driven by (1) specific information in sales forecasts and/or (2) the act of disclosing sales forecasts. I expect that the magnitude of sales forecast revision is positively associated with market reactions ($\beta_3 > 0$) and that supplementary sales forecasts add to the credibility of earnings forecasts ($\beta_2 > 0$). There should be stronger market reactions for the *same* amount of earnings forecast revisions when the revised sales forecasts are revealed.

Abarbanell et al. (1995) argue that inventors' response to earnings forecast revisions increases in the expected accuracy of the forecast. Because *ex post* accuracy of an earnings forecasts is unknown to the market when a financial analyst issues a forecast, investors use information available at the time of the forecast to estimate the *ex ante* perceived accuracy of the earnings forecast and react accordingly (Clement and Tse 2003). As a robustness check, I control for

[18] Sales forecasts could provide information about the persistence of future earnings because aggregate earnings mask the information on the persistence of different revenue and expense items by weighting them equally in the calculation of net income (Lambert 2004). Because I also find a positive coefficient on sales forecast revisions (β_3), it is likely that the persistence effect has been controlled for. I also partition my sample into five quintiles according to earnings growth, earnings change, and earnings persistence as measured in Beaver et al. (1980). My main results hold for firms from different quintiles of these firm characteristics (results not tabulated). Therefore, even if revisions in sales forecasts do not fully capture information about persistence, my results are not driven by the possibility that supplemented earnings forecast revisions are perceived as more persistent than stand-alone earnings forecast revisions.

individual analyst and revision characteristics that are associated with perceived accuracy, including forecast frequency (*FREQ*), forecast horizon (*HORIZON*), days elapsed since the last forecast (*DAYS*), prior earnings forecast error (*LAG_FE*), brokerage size (*BROKER*), firm-specific and general analyst experience (*FIRM_EXP* and *GEN_EXP*, respectively), the numbers of companies and industries followed by the analyst (*COMPANY* and *INDUSTRY*, respectively), and the level of earnings forecast revision innovation (*E_INNO*) (Clement 1999; Jacob et al. 1999; Brown 2001; Clement and Tse 2003; Gleason and Lee 2003).[19] Following Clement and Tse (2005) and Herrmann and Thomas (2005), I measure individual analyst characteristics of a financial analyst relative to the characteristics of other analysts following the same firm in the same year. I scale the individual analyst characteristic to lie between 0 and 1 by computing the difference between the observed value of the characteristic and the minimum value of the characteristic for the analysts following firm j in year t, divided by the range of the characteristic for the analysts following firm j in year t. The scaled individual analyst characteristic for a revision issued by analyst i for firm j in year t is given by:

$$SCALED\_CHAR_{ijt} = \frac{CHAR_{ijt} - \min_{jt}(CHAR)}{\max_{jt}(CHAR) - \min_{jt}(CHAR)} \tag{2}$$

where $\max_{jt}(CHAR)$ and $\min_{jt}(CHAR)$ are the maximum and the minimum values, respectively, of the individual analyst characteristic for the analysts following firm j in year t. This scaling preserves the relative distances among the individual analyst characteristics for firm j in year t and facilitates comparisons of regression coefficients across firms and over time.[20]

I run the following regression (subscripts omitted):

$$\begin{aligned} BHAR3 = {} & \beta_0 + \beta_1 E\_REV + \beta_2 S\_REV + \beta_3 E\_REV * DAYS + \beta_4 E\_REV * HORIZON \\ & + \beta_5 E\_REV * FREQ + \beta_6 E\_REV * LAG\_FE + \beta_7 E\_REV * NEW \\ & + \beta_8 E\_REV * BROKER + \beta_9 E\_REV * FIRM\_EXP + \beta_{10} E\_REV * GEN\_EXP \\ & + \beta_{11} E\_REV * COMPANY + \beta_{12} E\_REV * INDUSTRY + \beta_{13} E\_REV * E\_INNO \\ & + \beta_{14} E\_REV * SUPPL + \varepsilon \end{aligned} \tag{3}$$

[19] Another reason for including individual analyst characteristics is that supplementary sales forecasts are voluntary. If supplemented and stand-alone earnings forecast revisions are substantially different from one another, this would result in a self-selection bias. To access the severity of the self-selection bias, I used the Heckman (1979) approach to compute the inverse Mills ratio from the first-stage Probit model, with the individual analyst characteristics as first-stage independent variables. In the second-stage accuracy regression model (Equation (4)), the inverse Mills ratio is added as an additional control variable. The inverse Mills ratio is insignificant (coefficient estimate = 0.0276; p = 0.165). The estimate for regression coefficient on *SUPPL* is still negative and significant (coefficient estimate = −0.780; p = 0.031). Correcting for self-selection bias does not appear to be necessary. Moreover, the possibility of presence of self-selection in and of itself does not necessarily mean that OLS estimates are more biased than two-stage estimates (Wooldridge 2002). If the instruments are weak and/or partially endogenous, two-stage instrumental variable estimation can produce estimates that are more biased than OLS estimation. Two-stage estimates also have poor finite sample properties. Therefore, instead of including the inverse Mills ratios (Heckman 1979), I explicitly control for potential determinants of supplementary sales forecasts in the regressions to mitigate the endogeneity problem. Because not all determinants of disclosure of sales forecasts are observable to the econometrician, I could not totally rule out the possibility of self-selection. I caution the reader to keep this in mind in interpreting my results.

[20] I delete observations from firm-years with only one financial analyst issuing earnings forecasts because *relative* accuracy is not measurable.

where:

$BHAR3_{ijt}$ = three-day size-adjusted buy-and-hold return (−1 to +1) surrounding an annual earnings forecast revision. Size-adjusted buy-and-hold return is the raw buy-and-hold return of the firm minus the mean buy-and-hold return of an equal-weighted portfolio of firms in the same NYSE/AMEX/NASDAQ size decile over the holding period;

E_REV_{ijt} = revision in annual earnings forecast relative to the analyst's own prior forecast deflated by stock price two days before revision;

S_REV_{ijt} = revision in annual sales forecast relative to the analyst's own prior forecast deflated by stock price two days before revision if it is revealed, and 0 otherwise;

$DAYS_{ijt}$ = number of days elapsed since the last forecast by any analyst following firm j in year t, calculated as the number of days between analyst i's forecast of firm j's earnings in year t and the most recent preceding forecast of firm j's earnings by any analyst, minus the minimum number of days between two adjacent forecasts of firm j's earnings by any two analysts in year t, scaled by the range of days between two adjacent forecasts of firm j's earnings in year t;

$HORIZON_{ijt}$ = forecast horizon, calculated as the number of days from the forecast date to the fiscal year-end, for analyst i following firm j in year t minus the minimum forecast horizon for analysts who follow firm j in year t, scaled by the range of forecast horizons for analysts following firm j in year t;

$FREQ_{ijt}$ = forecast frequency, calculated as the number of forecasts made by analyst i for firm j in year t minus the minimum number forecasts made for firm j in year t out of all analysts following firm j, scaled by the range of the number of forecasts made for firm j issued by analysts following firm j in year t;

LAG_FE_{ijt} = absolute forecast error of last earnings forecast of analyst i following firm j in year $t-1$ minus the minimum absolute forecast error for analysts who follow firm j in year $t-1$, scaled by the range of absolute forecast error for analysts following firm j in year $t-1$;

NEW_{ijt} = 1 if the analyst i first appears in the I/B/E/S detail file in year t, and 0 otherwise;

$BROKER_{ijt}$ = the number of analysts employed by the broker employing analyst i following firm j in year t minus the minimum number of analysts employed by brokers for analysts following firm j in year t, scaled by the range of brokerage size for analysts following firm j in year t;

$FIRM\_EXP_{ijt}$ = firm-specific experience, which is the number of years analyst i has been following firm j in year t minus the minimum number of years of firm-specific experience for analysts following firm j in year t, scaled by the range of years of firm-specific experience for analysts following firm j in year t;

GEN_EXP_{ijt} = general experience, which is the number of years analyst i has been identified in the I/B/E/S detail file in year t minus the minimum number of years of general experience for analysts following firm j in year t, scaled by the range of years of general experience for analysts following firm j in year t;

$COMPANY_{ijt}$ = the number of companies analyst i follows in year t, calculated as the number of companies followed by analyst i following firm j in year t minus the minimum number of companies followed by analysts who follow firm j in

year t, scaled by the range of the number of companies followed by analysts following firm j in year t;

$INDUSTRY_{ijt}$ = the number of industries analyst i follows in year t, calculated as the number of two-digit SICs followed by analyst i following firm j in year t minus the minimum number of two-digit SICs followed by analysts who follow firm j in year t, scaled by the range of the number of two-digit SICs followed by analysts following firm j in year t;

E_INNO_{ijt} = 1 if the revised earnings forecast is either higher or lower than both analyst's own prior forecast and the prior day's consensus forecast (high innovation), and 0 otherwise; and

$SUPPL_{ijt}$ = 1 if the earnings forecast revision is supplemented by a sales forecast revision issued by the same analyst on the same day, and 0 otherwise.

In Equation (3), the coefficient on the interaction term (β_{14}) captures the incremental effect on the market reactions to supplemented earnings forecasts. The variable *SUPPL* equals 1 if the revision in earnings forecast is supplemented with sales forecast revision, and 0 otherwise. I expect that supplementary sales forecasts increase the credibility of earnings forecast revisions and lead to stronger market reactions ($\beta_{14} > 0$), controlling for the information content of sales forecasts ($\beta_2 > 0$). Individual analyst characteristics that are associated with increased forecast accuracy (e.g., prior accuracy, brokerage size, forecast frequency, and firm-specific and general experience) should be positively associated with market reactions to forecast revisions, while characteristics associated with decreased forecast accuracy (e.g., the numbers of companies and industries followed by the analyst) should be negatively associated with market reactions to forecast revisions.

Association between Contemporaneous Market Reaction and Supplementary Sales Forecast Revision

Table 2, Panel A reports the results of the regression analyses to examine the impacts of supplementary sales forecast revisions on stock prices. In Model I, the coefficient on earnings forecast revisions (β_1) is positive (0.4042) and significant ($p < 0.001$). This confirms the positive relationship between contemporaneous stock returns and earnings forecast revisions (Lys and Sohn 1990; Stickel 1991; Park and Stice 2000). The coefficient on sales forecast revision (β_3) is also positive and significant (0.0619; $p < 0.001$), controlling for revision in earnings forecast. Sales forecast revisions are incrementally informative to the market, which is consistent with the findings from prior research that sales are informative about stock returns (Ertimur et al. 2003; Ghosh et al. 2005).

In Model II, the coefficient on the interaction term between earnings forecast revision and the indicator variable for the presence of supplementary sales forecasts (β_2) is positive (0.2188) and significant ($p < 0.001$). The market reacts more strongly to supplemented earnings forecast revisions (0.4377) than to stand-alone earnings forecast revisions (0.2189), controlling for the independent information in sales forecast revisions. The reaction to earnings forecast revisions is twice as strong when accompanied by revenue forecast revisions. In Model III, I include both revision in sales forecasts and the interaction term in the regression. I find the coefficient on the interaction is still positive and significant (0.2158; $p < 0.001$), while the coefficient on sales forecast revision is marginally significant ($p = 0.069$). My findings support the conclusion that the market reaction is driven by *both* the specific information in sales forecast revisions and the act of disclosing sales forecasts. Supplementary sales forecast revisions appear to bolster the credibility of earnings forecasts.

TABLE 2

Effects of Supplementary Sales Forecast Revisions on the Market Reactions to Earnings Forecast Revisions

Panel A: Regression of Three-Day Buy-and-Hold Size-Adjusted Returns on Supplementary Sales Forecast Revisions

Model 1: $BHAR3 = \beta_0 + \beta_1 E\_REV + \beta_2 E\_REV * SUPPL + \beta_3 S\_REV + \varepsilon.$

Model 2: $BHAR3 = \beta_0 + \beta_1 E\_REV^{+} + \beta_2 E\_REV^{+} * SUPPL + \beta_3 E\_REV^{-} + \beta_4 E\_REV^{-} * SUPPL + \beta_5 S\_REV^{+} + \beta_6 S\_REV^{-} + \varepsilon.$

Dependent Variable = *BHAR3*

	I		II		III		IV	
Variable	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Intercept	−0.0020	<0.001	−0.0007	<0.001	−0.0007	<0.001	−0.0017	<0.001
E_REV	0.4042	<0.001	0.2189	<0.001	0.2189	<0.001		
*E_REV * SUPPL*			0.2188	<0.001	0.2158	<0.001		
S_REV	0.0619	<0.001			0.0017	0.069		
E_REV^{+}							0.5322	<0.001
$E\_REV^{+} * SUPPL$							0.1620	0.010
E_REV							0.1771	<0.001
$E\_REV^{-} * SUPPL$							0.2235	<0.001
S_REV^{+}							0.0046	0.010
S_REV^{-}							0.0006	0.574

(continued on next page)

Panel B: Regression of Three-Day Buy-and-Hold Size-Adjusted Returns on Supplementary Sales Forecast Revisions Controlling for Other Individual Analyst Characteristics

$$BHAR3_{ijt} = \beta_0 + \beta_1 E\_REV_{ijt} + \beta_2 S\_REV_{ijt} + \beta_3 E\_REV * DAYS_{ijt} + \beta_4 E\_REV * HORIZON_{ijt} + \beta_5 E\_REV * FREQ_{ijt} + \beta_6 E\_REV * LAG\_FE_{ijt} + \beta_7 E\_REV * NEW_{ijt} + \beta_8 E\_REV * BROKER_{ijt} + \beta_9 E\_REV * FIRM\_EXP_{ijt} + \beta_{10} E\_REV * GEN\_EXP_{ijt} + \beta_{11} E\_REV * COMPANY_{ijt} + \beta_{12} E\_REV * INDUSTRY_{ijt} + \beta_{13} E\_REV * E\_INNO_{ijt} + \beta_{14} E\_REV * SUPPL_{ijt} + \varepsilon_{ijt}.$$

Dependent Variable = *BHAR3*

	V		VI	
	Coefficient	**p-value**	**Coefficient**	**p-value**
Intercept	−0.0017	<0.001	−0.0017	<0.001
E_REV	0.4503	<0.001	0.4689	<0.001
S_REV	0.1174	<0.001	0.1335	<0.001
*E_REV * DAYS*	0.1584	0.000	0.1022	0.015
*E_REV * HORIZON*	−0.1269	0.009	−0.1117	0.016
*E_REV * FREQ*	0.3428	<0.001	0.3288	<0.001
*E_REV * LAG_FE*	−0.1224	0.000		
*E_REV * NEW*			0.2433	<0.001
*E_REV * BROKER*	0.2414	<0.001	0.2412	<0.001
*E_REV * FIRM_EXP*	−0.0393	0.331	−0.0164	0.685
*E_REV * GEN_EXP*	−0.0257	0.514	−0.0371	0.312
*E_REV * N_COMP*	−0.3553	0.003	−0.3051	0.007
*E_REV * N_IND*	0.3240	0.006	0.2771	0.014
*E_REV * E_INNO*	−0.0255	0.672	−0.1147	0.040
*E_REV * SUPPL*	0.3254	<0.001	0.3635	<0.001

This table presents the effects of supplementary sales forecast revisions on the magnitudes of forecast response coefficients relating the stock price responses to individual analyst earnings forecast revisions. Estimates for regression coefficients and two-tailed p-values for t-statistics are reported.

In Panel A, different coefficients on upward and downward earnings and sales forecast revisions are allowed.

In Panel B, I randomly select an earnings forecast revision by an analyst if s/he issues multiple revisions in a firm-year. I scale each of the individual analyst characteristics to lie between 0 and 1 by computing the difference between the observed value of the characteristic and the minimum value of the characteristic for the analysts following firm *j* in year *t*, divided by the range of the characteristic for the analysts following firm *j* in year *t*.

Variable Definitions:

$BHAR3_{ijt}$ = three-day size-adjusted buy-and-hold return (−1 to +1) around the annual earnings forecast revision by analyst *i* for firm *j* in year *t*;

E_REV_{ijt} = the revision in earnings forecast relative to the analyst *i*'s own prior forecast for firm *j* in year *t* deflated by stock price two days before revision;

(continued on next page)

$E\_REV^{+}_{ijt}$ = positive measure of revision in earnings forecast, i.e., $\max(E\_REV_{ijt}, 0)$;
$E\_REV^{-}_{ijt}$ = negative measure of revision in earnings forecast, i.e., $\min(E\_REV_{ijt}, 0)$;
S_REV_{ijt} = the revision in sales forecast relative to the analyst i's own prior forecast for firm j in year t deflated by stock price two days before revision;
$S\_REV^{+}_{ijt}$ = positive measure of revision in sales forecast, i.e., $\max(S\_REV_{ijt}, 0)$;
$S\_REV^{-}_{ijt}$ = negative measure of revision in sales forecast, i.e., $\min(S\_REV_{ijt}, 0)$;
$SUPPL_{ijt}$ = 1 if there is a supplementary sales forecast revision on the same day as an earnings forecast revision by the same analyst for the same firm, and 0 otherwise;
FE_{ijt} = absolute forecast error for analyst i who follows firm j in year t minus the minimum absolute forecast error for the analysts following firm j in year t, scaled by the range of absolute forecast error for the analysts following firm j in year t;
$DAYS_{ijt}$ = number of days elapsed since the last forecast by any analyst following firm j in year t, calculated as the number of days between analyst i's forecast of firm j's earnings in year t and the most recent preceding forecast of firm j's earnings by any analyst, minus the minimum number of days between two adjacent forecasts of firm j's earnings by any two analysts in year t, scaled by the range of days between two adjacent forecasts of firm j's earnings in year t;
$HORIZON_{ijt}$ = forecast horizon, calculated as the number of days from the forecast date to the fiscal year-end, for analyst i following firm j in year t minus the minimum forecast horizon for analysts who follow firm j in year t, scaled by the range of forecast horizons for analysts following firm j in year t;
$FREQ_{ijt}$ = forecast frequency, calculated as the number of forecasts made by analyst i for firm j in year t minus the minimum number forecasts made for firm j in year t out of all analysts following firm j, scaled by the range of the number of forecasts made for firm j issued by analysts following firm j in year t;
LAG_FE_{ijt} = absolute forecast error of last earnings forecast of analyst i following firm j in year $t-1$ minus the minimum absolute forecast error for analysts who follow firm j in year $t-1$, scaled by the range of absolute forecast error for analysts following firm j in year $t-1$;
NEW_{ijt} = 1 if the analyst i first appears in the I/B/E/S detail file in year t, and 0 otherwise;
$BROKER_{ijt}$ = the number of analysts employed by the brokerage employing analyst i following firm j in year t minus the minimum number of analysts employed by brokerages for analysts following firm j in year t, scaled by the range of brokerage size for analysts following firm j in year t;
$FIRM\_EXP_{ijt}$ = firm-specific experience, which is the number of years analyst i has been following firm j in year t minus the minimum number of years of firm-specific experience for analysts following firm j in year t, scaled by the range of years of firm-specific experience for analysts following firm j in year t;
GEN_EXP_{ijt} = general experience, which is the number of years analyst i has been identified in the I/B/E/S detail file in year t minus the minimum number of years of general experience for analysts following firm j in year t, scaled by the range of years of general experience for analysts following firm j in year t;
$COMPANY_{ijt}$ = the number of companies analyst i follows in year t, calculated as the number of companies followed by analyst i following firm j in year t minus the minimum number of companies followed by analysts who follow firm j in year t, scaled by the range of the number of companies followed by analysts following firm j in year t;
$INDUSTRY_{ijt}$ = the number of industries analyst i follows in year t, calculated as the number of two-digit SICs followed by analyst i following firm j in year t minus the minimum number of two-digit SICs followed by analysts who follow firm j in year t, scaled by the range of the number of two-digit SICs followed by analysts following firm j in year t; and
E_INNO_{ijt} = 1 for high-innovation earnings forecast revisions, and 0 for low-innovation earnings forecast revisions.

I further investigate whether the effect of a supplementary sales forecast revision is conditional on the direction of the earnings forecast revision. Each earnings forecast revision is classified as an upward or a downward revision based on whether the revised forecast is above or below the prior forecast. Results in Model IV show that the market reacts more strongly to upward earnings forecast revisions (0.5322) than to downward revisions (0.1771). However, the presence of supplementary sales forecasts increases the market response to downward earnings forecast revisions (0.2235) to a greater extent than it increases the response to upward earnings forecast revisions (0.1620). There is no market reaction to the independent information in downward sales forecast revisions after controlling for earnings forecast revision.

Table 2, Panel B reports additional results on the impact of supplementary sales forecast revisions on stock prices. I control for other individual analyst characteristics, which could be potential determinants of supplementary sales forecasts. In Model V, I control for prior forecast accuracy (*LAG_FE*). I measure prior earnings forecast accuracy using the absolute forecast error of the last earnings forecasts provided by the analyst.[21] The coefficients on earnings and sales forecast revisions (β_1 and β_2, respectively) remain positive and significant. The coefficient on the interaction between earnings forecast revision and the presence of supplementary sales forecasts (β_{14}) is positive (0.3254) and significant ($p < 0.001$), controlling for other individual analyst and forecast revision characteristics. The coefficients on the control variables are consistent with the finding in Clement and Tse (2003) that market reactions increase in *ex ante* perceived accuracy. For example, days since last forecasts (*DAYS*), forecast frequency (*FREQ*), brokerage size (*BROKER*), and number of industries followed (*INDUSTRY*) are positively associated with market reactions to earnings forecast revisions, while forecast horizon (*HORIZON*) and number of companies followed (*COMPANY*) are negatively associated with market reactions to earnings forecast revisions.

In Model VI, I retain observations from financial analysts that first appear in the I/B/E/S database in the year by dropping prior forecast accuracy (*LAG_FE*). Instead, I include the indicator variable *NEW*, where *NEW* equals 1 if the financial analyst first appears in the I/B/E/S database in that year, and 0 otherwise. The results are similar to those from Model V. In summary, the results in Table 2 suggest that the market reacts more strongly to supplemented earnings forecasts. Supplementary sales forecasts add to the credibility of earnings forecast revisions, controlling for independent information in sales forecasts and other individual analyst and revision characteristics. In Section V below I explore the causes of the stronger market reactions by investigating whether supplemented earnings forecasts are more accurate *ex post*.

V. ASSOCIATION BETWEEN PROPERTIES OF EARNINGS FORECAST REVISIONS AND SUPPLEMENTARY SALES FORECASTS

Research Design

Following prior literature, I model forecast accuracy as a function of forecast frequency (*FREQ*), forecast horizon (*HORIZON*), days elapsed since the last forecast (*DAYS*), prior earnings forecast accuracy (*LAG_FE*), brokerage size (*BROKER*), analyst experience (*FIRM_EXP* and *GEN_EXP*), and the numbers of companies and industries followed by the analyst (*COMPANY* and *INDUSTRY*) (Clement 1999; Jacob et al. 1999; Brown 2001; Clement and Tse 2003). I also include the magnitudes of earnings and sales forecast revisions (|*E_REV*| and |*S_REV*|) and the

[21] Absolute forecast error of last earnings forecast provided by the analyst might be a noisy measure for forecast accuracy of financial analysts because last forecasts in fiscal years are issued at different points in time by different financial analysts.

level of earnings forecast revision innovation (*E_INNO*) to control for the information content of the earnings and the sales forecasts. I run the following regression:

$$\begin{aligned}|FE_{ijt}| = {} & \beta_0 + \beta_1 DAYS_{ijt} + \beta_2 HORIZON_{ijt} + \beta_3 FREQ_{ijt} + \beta_4 LAG\_FE_{ijt} + \beta_5 NEW_{ijt} \\ & + \beta_6 BROKER_{ijt} + \beta_7 FIRM\_EXP_{ijt} + \beta_8 GEN\_EXP_{ijt} + \beta_9 COMPANY_{ijt} \\ & + \beta_{10} INDUSTRY_{ijt} + \beta_{11} E\_INNO_{ijt} + \beta_{12} |E\_REV_{ijt}| + \beta_{13} |S\_REV_{ijt}| + \beta_{14} SUPPL_{ijt} \\ & + \varepsilon_{ijt} \end{aligned} \quad (4)$$

where:

$|FE_{ijt}|$ = absolute value of forecast error for analyst i who follows firm j in year t minus the minimum absolute forecast error for the analysts following firm j in year t, scaled by the range of absolute forecast error for the analysts following firm j in year t.

Definitions of other independent variables are the same as in previous analyses. The test variable, *SUPPL*, is an indicator variable for the presence of a supplementary sales forecast. *SUPPL* equals 1 if the revision in earnings forecast is supplemented with a sales forecast revision, and 0 otherwise. I expect that forecast accuracy increases with prior accuracy (*LAG_FE*), brokerage size (*BROKER*), forecast frequency (*FREQ*), analyst experience (*FIRM_EXP* and *GEN_EXP*), and the level of revision innovation (*E_INNO*), but decreases with days elapsed since the last forecast (*DAYS*), forecast horizon (*HORIZON*), and the numbers of companies and industries followed by the analyst (*COMPANY* and *INDUSTRY*, respectively)[22] If earnings forecasts supplemented with sales forecast revisions are more accurate, then the coefficient on *SUPPL* (β_{14}) should be negative (H2). A negative coefficient means that supplementary sales forecasts are associated with earnings forecasts with lower absolute forecast errors. If supplemented earnings forecasts are issued by the less informed financial analysts, then the presence of supplementary sales forecasts should not be associated with earnings forecast accuracy.

Univariate Analysis of Supplemented and Stand-Alone Earnings Forecast Revisions

I compare the means of absolute forecast error and other individual analyst characteristics for supplemented and stand-alone earnings forecast revisions in Table 3, Panel A. Most of the individual analyst characteristics differ significantly across supplemented and stand-alone earnings forecast revisions under both two-sample t-tests and nonparametric Wilcoxon rank-sum tests.[23] The means for days elapsed since the last forecast (*DAYS*), forecast frequency (*FREQ*), number of firms followed (*COMPANY*), and number of industries followed (*INDUSTRY*) are higher for supplemented earnings forecast revisions, whereas the means for absolute forecast error (|*FE*|), forecast horizons (*HORIZON*), brokerage size (*BROKER*), firm-specific experience (*FRIM_EXP*), and general experience (*GEN_EXP*) are lower for supplemented earnings forecast revisions. Al-

[22] I include both firm-specific and general experience because I argue that, *a priori*, they are proxies for different dimensions of analyst attributes. General experience could proxy for an analyst's general industry knowledge, while firm-specific experience could proxy for an analyst's familiarity with a specific firm. For the same reason I include both numbers of companies and industries followed by the analyst. Empirically, firm-specific experience and general experience are highly correlated. Nevertheless, my results are robust to including both of them or dropping either one of them from the regressions.

[23] The parametric two-sample t-test tests the null hypothesis that the mean difference of two populations is zero, assuming that the observations are drawn from normal distributions with equal variances. The nonparametric Wilcoxon rank sum test compares the locations of two populations to determine if one population is shifted with respect to the other.

TABLE 3

Descriptive Statistics on Individual Analyst Characteristics

Panel A: Comparison of Individual Analyst Characteristics for Supplemented and Stand-Alone Earnings Forecast Revisions

	Supplemented Revisions		Stand-Alone Revisions		Two-Sample t-test		Wilcoxon Rank-Sum Test	
Variables	**Mean**	**Median**	**Mean**	**Median**	**t-stat**	**p-value**	**z-score**	**p-value**
|FE|	0.3427	0.2361	0.3713	0.2834	10.32	<0.001	10.47	<0.001
DAYS	0.3180	0.1919	0.2943	0.1795	−9.07	<0.001	−5.87	<0.001
HORIZONS	0.4668	0.4511	0.4998	0.5000	12.09	<0.001	11.91	<0.001
FREQ	0.5725	0.5714	0.5287	0.5000	−15.34	<0.001	−15.56	<0.001
LAG_FE	0.4124	0.5000	0.4056	0.4286	−2.05	0.041	−2.09	0.036
BROKER	0.3500	0.2354	0.3677	0.2613	5.95	<0.001	1.10	0.271
FIRM_EXP	0.4256	0.3333	0.4781	0.5000	16.36	<0.001	16.22	<0.001
GEN_EXP	0.5066	0.5000	0.6062	0.6667	30.78	<0.001	31.45	<0.001
N_COMP	0.4839	0.4545	0.4465	0.4000	−13.12	<0.001	−13.15	<0.001
N_IND	0.4793	0.4444	0.4411	0.3871	−13.37	<0.001	−13.62	<0.001
E_INNO	0.7246	1.0000	0.7325	1.0000	2.13	0.034	2.13	0.034
|E_REV|	0.0125	0.0027	0.0134	0.0030	1.92	0.055	5.35	<0.001
n	15,786		116,360					

Panel B: Pearson Correlations Among Individual Analyst Characteristics (p-values reported below correlations)

	|FE|	*DAYS*	*HORIZONS*	*FREQ*	*LAG_FE*	*BROKER*	*FIRM_EXP*
DAYS	0.0307						
	<0.001						
HORIZONS	0.458	−0.0439					
	<0.001	<0.001					
FREQ	0.0196	0.0459	−0.041				
	<0.001	<0.001	<0.001				
LAG_FE	−0.0055	0.0026	−0.0041	0.0047			
	0.064	0.374	0.164	0.115			
BROKER	0.0068	0.0691	−0.0008	0.056	0.0018		
	0.014	<0.001	0.760	<0.001	0.553		

(continued on next page)

Panel B: Pearson Correlations Among Individual Analyst Characteristics (p-values reported below correlations)

	|FE|	*DAYS*	*HORIZONS*	*FREQ*	*LAG_FE*	*BROKER*	*FIRM_EXP*
FIRM_EXP	0.0863	0.0415	0.1622	0.0569	0.0018	0.0681	
	<0.001	<0.001	<0.001	<0.001	0.539	<0.001	
GEN_EXP	0.0434	−0.0097	0.0931	0.0094	0.0068	0.0601	0.5071
	<0.001	0.000	<0.001	0.001	0.023	<0.001	<0.001
COMPANY	0.0169	0.0311	0.0153	0.089	0.0052	0.0815	0.1844
	<0.001	<0.001	<0.001	<0.001	0.081	<0.001	<0.001
INDUSTRY	0.0166	0.031	0.0148	0.0858	0.0046	0.0887	0.1756
	<0.001	<0.001	<0.001	<0.001	0.125	<0.001	<0.001
E_INNO	−0.0009	−0.0042	0.0237	−0.0323	0.005	0.0153	0.029
	0.745	0.121	<0.001	<0.001	0.093	<0.001	<0.001
|E_REV|	−0.0221	0.0163	0.0046	−0.0432	0.0008	0.0221	0.023
	<0.001	<0.001	0.096	<0.001	0.800	<0.001	<0.001
NEW	−0.1333	−0.0026	−0.2842	−0.0168	0.001	−0.0009	−0.5377
	<0.001	0.346	<0.001	<0.001	0.746	0.742	<0.001
SUPPL	−0.0309	0.0255	−0.0389	0.0424	0.0011	−0.0178	−0.0456
	<0.001	<0.001	<0.001	<0.001	0.718	<0.001	<0.001

	GEN_EXP	*COMPANY*	*INDUSTRY*	*E_INNO*	*|E_REV|*	*NEW*
COMPANY	0.2913					
	<0.001					
INDUSTRY	0.2842	0.9582				
	<0.001	<0.001				
E_INNO	0.0155	0.0124	0.014			
	<0.001	<0.001	<0.001			
|E_REV|	−0.0062	0.0152	0.0142	0.0955		
	0.025	<0.001	<0.001	<0.001		
NEW	−0.3097	−0.0985	−0.0924	−0.018	−0.0247	
	<0.001	<0.001	<0.001	<0.001	<0.001	
SUPPL	−0.0849	0.0362	0.0368	−0.0058	−0.0068	0.0399
	<0.001	<0.001	<0.001	0.034	0.013	<0.001

(continued on next page)

This table reports descriptive statistics on the individual analyst characteristics. Individual analyst characteristics are obtained from the I/B/E/S detail file. I scale each of the individual analyst characteristics to lie between 0 and 1 by computing the difference between the observed value of the characteristic and the minimum value of the characteristic for the analysts following firm j in year t, divided by the range of the characteristic for the analysts following firm j in year t. Data on forecast accuracy for prior year is available for 96,481 observations. Supplementary sales forecasts are available for 15,786 revisions.

Panel A compares individual analyst characteristics for supplemented and stand-alone earnings forecast revisions. Corresponding t-statistics and z-scores and their p-values (two-tailed) are reported.

Panel B reports Pearson correlations among the individual analyst characteristics.

See Table 2 for variable definitions.

though financial analysts who provide supplementary sales forecasts are more likely to be less experienced and employed by smaller brokerage houses, supplemented earnings forecast revisions are more accurate than stand-alone earnings forecast revisions.

Table 3, Panel B reports Pearson correlations among individual analyst characteristics. The presence of a supplementary sales forecast revision (*SUPPL*) is positively correlated with forecast accuracy (negative correlation with absolute forecast error, |*FE*|). Forecast accuracy is negatively correlated with days elapsed since the last forecast (*DAYS*), forecast horizon (*HORIZON*), and the numbers of companies and industries followed by the financial analyst (*COMPANY* and *INDUSTRY*). Prior accuracy is not significantly correlated with the accuracy of earnings forecast in the current year. One possible reason is that prior research measures forecast accuracy using the *last* earnings forecast issued by a financial analyst in the previous year. In contrast, I randomly pick an earnings forecast issued *anytime* during the current year. If a more able financial analyst issues earnings forecasts earlier in the year at the cost of lower accuracy, then the accuracy of his/her earlier forecasts in the current year might not be related to the accuracy of his/her *last* forecast in the prior year.

Association between Earnings Forecast Accuracy and Supplementary Sales Forecasts

Table 4 reports regression results for the association between the magnitude of earnings forecast error and the presence of a supplementary sales forecast. In Model I, the coefficient on the supplementary sales forecast indicator variable (*SUPPL*) is negative. This supports the prediction (H2) that earnings forecasts supplemented with sales forecast revisions are more accurate, controlling for other individual analyst characteristics. Although the differences in forecast accuracy seem to be economically small, as shown in Table 2, market reactions are twice as strong for supplemented earnings forecast revisions. An apparently small increment in *ex post* forecast accuracy could trigger a much stronger *ex ante* market reaction. Note that the coefficient on *SUPPL* picks up its partial effect on forecast accuracy, helping to explain why the magnitude of the coefficient seems to be low. Consistent with prior literature, I also find that forecast accuracy increases with general experience and decreases with days lapsed since the last forecast, forecast horizon, and forecast frequency. High-innovation earnings forecasts are, on average, more accurate than low-innovation earnings forecasts.

In Model II, the indicator variable *NEW* equals 1 if it is the first year that the financial analyst appears in the I/B/E/S detail file, and 0 otherwise. The results suggest that earnings forecasts with supplementary sales forecast revisions are more accurate even if I include earnings forecasts provided by new financial analysts in the sample. In Models III and IV, I include |*S_REV*| as an additional independent variable. The coefficient on |*S_REV*| is insignificant. This means that earnings forecast accuracy is not driven mechanically by the magnitude of sales forecast revision.

As a robustness check, I repeat my analysis using the last earnings forecast by an analyst in each firm-year (Sinha et al. 1997; Clement and Tse 2005; Herrmann and Thomas 2005). The results are qualitatively the same and are not tabulated. Furthermore, I estimate Equation (4) using the *change* in absolute forecast error as the dependent variable (Clement and Tse 2005). Consistent with using the *level* of absolute forecast error as the dependent variable in Table 4, untabulated results show that the regression coefficient on the indicator variable for the presence of supplementary sales forecasts (*SUPPL*) is negative and statistically significant, suggesting that supplemented earnings forecast revisions lead to a greater decrease in the magnitude of forecast errors.

TABLE 4

Association between Forecast Accuracy and Supplementary Sales Forecast Revisions

$$|FE_{ijt}| = \beta_0 + \beta_1 DAYS_{ijt} + \beta_2 HORIZON_{ijt} + \beta_3 FREQ_{ijt} + \beta_4 LAG\_FE_{ijt} + \beta_5 NEW_{ijt} + \beta_6 BROKER_{ijt} + \beta_7 FIRM\_EXP_{ijt} + \beta_8 GEN\_EXP_{ijt} + \beta_9 COMPANY_{ijt} + \beta_{10} INDUSTRY_{ijt} + \beta_{11} E\_INNO_{ijt} + \beta_{12} |E\_REV_{ijt}| + \beta_{13} |S\_REV_{ijt}| + \beta_{14} SUPPL_{ijt} + \varepsilon_{ijt}.$$

Dependent Variable = $|FE|$

	I		II		III		IV	
	Coefficient	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Intercept	0.1274	<0.001	0.1121	<0.001	0.0736	<0.001	0.0721	<0.001
DAYS	0.0564	<0.001	0.0520	<0.001	0.0507	<0.001	0.0519	<0.001
HORIZON	0.4560	<0.001	0.4577	<0.001	0.4417	<0.001	0.4409	<0.001
FREQ	0.0283	<0.001	0.0333	<0.001	0.0386	<0.001	0.0457	<0.001
LAG_FE	0.0000	0.154			0.0341	<0.001		
NEW			0.0022	0.335			−0.0010	0.613
BROKER	−0.0011	0.660	0.0012	0.605	−0.0024	0.303	−0.0009	0.681
FIRM_EXP	−0.0007	0.799	0.0102	0.000	0.0032	0.217	0.0092	0.000
GEN_EXP	−0.0090	0.001	−0.0067	0.007	−0.0054	0.034	−0.0022	0.326
N_COMP	0.0025	0.776	0.0019	0.818	−0.0138	0.125	−0.0072	0.375
N_IND	0.0052	0.562	0.0047	0.567	0.0202	0.024	0.0125	0.123
E_INNO	−0.0093	<0.001	−0.0064	0.000	0.0003	0.845	0.0034	0.030
|E_REV|	−0.1577	<0.001	−0.1443	<0.001	−1.1242	<0.001	−1.0009	<0.001
|S_REV|					0.2512	0.408	0.2455	0.467
SUPPL	−0.0184	<0.001	−0.0165	<0.001	−0.0366	<0.001	−0.0340	<0.001
F-value	2470.55	<0.001	3041.37	<0.001	1832.66	<0.001	2396.01	<0.001
Adj. R^2	0.2073		0.2146		0.165		0.1741	

This table reports the association between forecast accuracy and supplementary sales forecast revisions, controlling for other individual analyst characteristics. I report the coefficient estimates and the p-values for t-statistics from pooled OLS regression.

See Table 2 for variable definitions.

VI. ADDITIONAL ANALYSES

Initiation and Termination of Sales Forecasts

In this section I investigate whether earnings forecasts are more or less accurate when financial analysts start or stop providing sales forecasts.[24] I define two additional indicator variables *START* and *STOP* and estimate the following regression:

$$\begin{aligned}|FE_{ijt}| = {} & \beta_0 + \beta_1 DAYS_{ijt} + \beta_2 HORIZON_{ijt} + \beta_3 FREQ_{ijt} + \beta_4 LAG\_FE_{ijt} + \beta_5 NEW_{ijt} \\ & + \beta_6 BROKER_{ijt} + \beta_7 FIRM\_EXP_{ijt} + \beta_8 GEN\_EXP_{ijt} + \beta_9 COMPANY_{ijt} \\ & + \beta_{10} INDUSTRY_{ijt} + \beta_{11} E\_INNO_{ijt} + \beta_{12} |E\_REV_{ijt}| + \beta_{13} SUPPL_{ijt} + \beta_{14} START_{ijt} \\ & + \beta_{15} STOP_{ijt} + \varepsilon_{ijt} \end{aligned} \quad (5)$$

where:

$|FE_{ijt}|$ = absolute value of forecast error for analyst i who follows firm j in year t minus the minimum absolute forecast error for the analysts following firm j in year t, scaled by the range of absolute forecast error for the analysts following firm j in year t;

$START_{ijt}$ = 1 if analyst i issues at least one sales forecast for firm j in year t but does not do so in year $t-1$, and 0 otherwise; and

$STOP_{ijt}$ = 1 if analyst i issues at least one sales forecast for firm j in year $t-1$ but does not do so in year t, and 0 otherwise.

As in Equation (4), the dependent variable is the absolute value of forecast error. The test variables are *START* and *STOP*. Definitions of other independent variables are the same as in previous analyses. A negative (positive) coefficient on *START* or *STOP* means that earnings forecasts are more (less) accurate when financial analysts initiate or terminate supplementary sales forecasts, respectively.

I identify 24,972 cases of initiation of sales forecasts (*START*) and 3,614 cases of termination (*STOP*) at the firm-year-analyst level. Table 5 reports the results. Consistent with results in Table 4, the coefficients on *SUPPL* are negative in all three models. In other words, the presence of supplementary sales forecasts improves forecast accuracy, controlling for *START* and *STOP*. After including other individual analyst characteristics in Models II and III, the *sums* of the coefficients on *SUPPL* and *START* are negative (−0.0221 and −0.0220 for Models II and III, respectively). The presence of supplementary sales forecasts plays a signaling role even for the initial issuance of sales forecasts. This also suggests that there is further improvement in earnings forecast accuracy beyond the year of initiation of sales forecasts.

The coefficients on *STOP* are positive but insignificant in all three models. One possible reason is that there are relatively few cases (less than 3 percent of total observations) of termination of sales forecasts. On the whole, these results support the argument there is an improvement in forecast accuracy when financial analysts start issuing sales forecasts.

Earnings Persistence and Firm Characteristics

To further assess the effects of earnings persistence on supplementary sales forecasts, I estimate the following regression:

[24] I use a cross-sectional regression approach to examine the changes in earnings forecast accuracy at the firm-analyst level around the initiation and termination of sales forecasts. I do not adopt a time-series research design because restricting the analyses to analysts that have survived for a certain period of time may induce results that are sensitive to survivor bias.

TABLE 5

Association between Forecast Accuracy and Initiation/Termination of Supplementary Sales Forecasts

$$|FE_{ijt}| = \beta_0 + \beta_1 DAYS_{ijt} + \beta_2 HORIZON_{ijt} + \beta_3 FREQ_{ijt} + \beta_4 LAG\_FE_{ijt} + \beta_5 NEW_{ijt} + \beta_6 BROKER_{ijt} + \beta_7 FIRM\_EXP_{ijt} + \beta_8 GEN\_EXP_{ijt} + \beta_9 COMPANY_{ijt} + \beta_{10} INDUSTRY_{ijt} + \beta_{11} E\_INNO_{ijt} + \beta_{12} |E\_REV_{ijt}| + \beta_{13} SUPPL_{ijt} + \beta_{14} START_{ijt} + \beta_{15} STOP_{ijt} + \varepsilon_{ijt}$$

Dependent Variable = $|FE|$

	I		II		III	
	Coefficient	**p-value**	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Intercept	0.3251	<0.001	0.0725	<0.001	0.0720	<0.001
DAYS			0.0494	<0.001	0.0508	<0.001
HORIZON			0.4435	<0.001	0.4422	<0.001
FREQ			0.0380	<0.001	0.0454	<0.001
LAG_FE			0.0347	<0.001		
NEW					−0.0018	0.394
BROKER			−0.0026	0.282	−0.0011	0.598
FIRM_EXP			0.0045	0.090	0.0103	<0.001
GEN_EXP			−0.0050	0.055	−0.0021	0.362
N_COMP			−0.0143	0.117	−0.0076	0.358
N_IND			0.0200	0.027	0.0121	0.143
E_INNO			0.0006	0.750	0.0037	0.022
|E_REV|			−1.0810	<0.001	−0.9645	<0.001
SUPPL	−0.0095	0.002	−0.0332	<0.001	−0.0320	<0.001
START	−0.0518	<0.001	0.0111	0.012	0.0100	0.009
STOP	0.0511	0.532	0.0886	0.242	0.0890	0.232
F-value	150.01	<0.001	1613.26	<0.001	2111.58	<0.001
Adj. R^2	0.0032		0.1651		0.1741	

This table reports the association between forecast accuracy and initiation/termination of supplementary sales forecasts, controlling for other individual analyst characteristics. I report the coefficient estimates and the p-values for t-statistics from pooled OLS regression.

Variable Definitions:

$START_{ijt}$ = 1 if analyst i issues at least one sales forecast for firm j in year t but does not do so in year $t-1$, and 0 otherwise; and

$STOP_{ijt}$ = 1 if analyst i issues at least one sales forecast for firm j in year $t-1$ but does not do so in year t, and 0 otherwise.

See Table 2 for definitions of other variables.

$$EPS_{t+1} = \beta_0 + \beta_1 EPS_t + \beta_2 SUPPL_t * EPS_t + \varepsilon_{t+1} \quad (6)$$

where EPS_{t+1} is earnings per share before extraordinary items. I estimate the above regression with Huber-White standard errors clustered by firm. Untabulated results show that the coefficient on EPS_t (β_1) is positively significant (coefficient estimate = 0.7841; $p < 0.001$), but the coefficient (β_2) on $SUPPL_t * EPS_t$ is insignificant (coefficient estimate = 0.0012; $p = 0.402$). My results are similar when including other individual analyst characteristics as control variables. Therefore, it is reasonable to conclude that the primary findings are robust to controlling for earnings persistence.

I also partition the sample into five quintiles according to each of the following firm characteristics: earnings persistence, earnings growth, earnings volatility, earnings change, firm size, and book-to-market ratio.[25] I run separate regressions for each quintile and find that the main results hold for firms from different quintiles of these firm characteristics (results not tabulated).

My sample consists of firm-years from firms that may not have sales forecasts. However, the reported conclusions are robust to restricting the sample to earnings forecast revisions from firm-years with at least one supplementary sales forecast revision. These results suggest that my main results are not driven by the possibility that sales forecasts might be more common for some particular firms.[26]

Time Effects

Although the reported analyses capture relative forecast accuracy and individual analyst characteristics *relative* to those of other financial analysts following the same firm in the same year, I assess the robustness of these results to including a time trend variable. The main results are qualitatively the same. Results are also robust to inclusion of year indicators for fixed time effects.

The Directions of Earnings Forecast Revisions and Sales Forecast Revisions

To provide further evidence on whether earnings forecast revisions and sales forecast revisions are in the same direction, I define two indicator variables, *CONFIRM* and *CONFLICT*. *CONFIRM* equals 1 if the revision in sales forecast is in the same direction as the revision in earnings forecasts, and 0 otherwise. Similarly, *CONFLICT* equals 1 if the revision in sales forecast is in the opposite direction as the revision in earnings forecasts, and 0 otherwise. I identify 15,786 cases of confirming sales forecasts (*CONFIRM*) and 4,146 cases of conflicting sales forecasts (*CONFIRM*) at the firm-year-analyst level. Instead of including *SUPPL* as the test variable in the regressions, I replace *SUPPL* with the variables *CONFIRM* and *CONFLICT*. Regression coefficients on both *CONFIRM* and *CONFLICT* are negative and significant. The main results (untabulated) are essentially the same, although the statistical significance on the *CONFLICT* variable is weaker. I also remove observations for which the revenue and the earnings forecast revisions are in the opposite directions. Again, the main findings are the same. I conclude that it is the *act* of disclosing sales forecasts, not the content of the sales forecasts that accounts for the primary results.

VII. CONCLUSION

Although a vast amount of academic research has examined different aspects of financial analysts' earnings forecasts, the interaction between earnings forecasts and other supplementary information provided by financial analysts is largely unexplored. In this study, I examine the nature and the market responses to financial analysts' supplementary sales forecasts. I find that earnings forecast revisions supplemented with sales forecast revisions have a greater impact on security prices, controlling for the incremental information content in sales forecasts. The differences in return responses suggest that investors view supplemented earnings forecast revisions as more credible, and that verifiable supplementary information increases the credibility of earnings forecasts. Furthermore, supplemented earnings forecasts are more accurate *ex post*, suggesting that financial analysts provide sales forecasts to convey their credibility. To the extent that the provi-

[25] Beaver et al. (1980) characterize the time-series of annual earnings as a first-order moving average process in the first differences, IMA(1,1). Following their approach, I measured earnings persistence as the moving average parameter (θ) in the IMA(1,1) process.

[26] Note that this study differs from prior research (e.g., DeFond and Hung 2003), as I use relative forecast accuracy as dependent variable and carry out the analyses at the analyst, not at the firm level.

sion of sales forecasts is associated with analyst ability, it is an open question as to whether less informed financial analysts would be penalized if they issue *ex post* inaccurate sales forecasts to mimic their more informed counterparts. Although sales forecasts are one possible way to lend credibility to earnings forecasts, there could also be other nonfinancial information disclosed in analyst reports that impacts the market reactions to earnings forecasts. I leave this to future research.

REFERENCES

Abarbanell, J., W. Lanen, and R. Verrecchia. 1995. Analysts' forecasts as proxies for investor beliefs in empirical research. *Journal of Accounting and Economics* 1 (20): 31–60.

Asquith, P., M. Mikhail, and A. Au. 2005. Information content of equity analyst reports. *Journal of Financial Economics* 75 (2): 245–282.

Baginski, S., J. Hassell, and M. Kimbrough. 2004. Why do managers explain their earnings forecasts? *Journal of Accounting Research* 42 (1): 1–29.

Beaver, W., R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting and Economics* 2 (1): 3–28.

Bradshaw, M. 2002. The use of target prices to justify sell-side analysts' stock recommendations. *Accounting Horizons* 16 (1): 27–41.

Brown, L. 2001. Predicting individual analyst earnings forecast accuracy. *Financial Analysts Journal* 57 (6): 44–49.

Call, A., S. Chen, and Y. Tong. 2009. Are analysts' earnings forecasts more accurate when accompanied by cash flow forecasts? *Review of Accounting Studies* 14 (2–3): 358–391.

Chen, Q., J. Francis, and W. Jiang. 2005. Investor learning about analyst predictive ability. *Journal of Accounting and Economics* 39 (1): 3–24.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27: 285–303.

———, and S. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78 (1): 227–249.

———, and ———. 2005. Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance* 60 (1): 307–341.

DeFond, M., and M. Hung. 2003. An empirical analysis of analysts' cash flow forecasts. *Journal of Accounting and Economics* 35 (1): 73–100.

Ertimur, Y., J. Livnat, and M. Martikainen. 2003. Differential market reactions to revenue and expense surprises. *Review of Accounting Studies* 8 (2): 185–211.

Francis, J., and L. Soffer. 1997. The relative informativeness of analysts' stock recommendations and earnings forecast revisions. *Journal of Accounting Research* 35 (2): 193–211.

Ghosh, A., Z. Gu, and P. Jain. 2005. Sustained earnings and revenue growth, earnings quality, and earnings response coefficients. *Review of Accounting Studies* 10 (1): 33–57.

Givoly, D., C. Hayn, and R. Lehavy. 2009. The quality of analysts' cash flow forecasts. *The Accounting Review* 84 (6): 1877–1911.

Gleason, C., and C. Lee. 2003. Analyst forecast revisions and market price discovery. *The Accounting Review* 78 (1): 193–225.

Hayes, R. 1998. The impact of trading commission incentives on analysts' stock coverage decisions and earnings forecasts. *Journal of Accounting Research* 36: 299–320.

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47 (1): 153–161.

Herrmann, D., and W. Thomas. 2005. Rounding of analyst forecasts. *The Accounting Review* 80 (3): 805–823.

Hutton, A., G. Miller, and D. Skinner. 2003. The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41 (5): 867–890.

Irvine, P. 2000. Do analysts generate trade for their firms? Evidence from the Toronto stock exchange. *Journal of Accounting and Economics* 30 (2): 209–226.

———. 2004. Analysts' forecasts and brokerage-firm trading. *The Accounting Review* 79 (1): 125–149.

Jackson, A. 2005. Trade generation, reputation, and sell-side analysts. *The Journal of Finance* 60 (2): 673–717.

Jacob, J., T. Lys, and M. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28: 27–50.

Lambert, R. 2004. Discussion of: Analysts' treatment of non-recurring items in street earnings and loss function assumptions in rational expectations tests on financial analysts' earnings forecasts. *Journal of Accounting and Economics* 38: 205–222.

Lys, T., and S. Sohn. 1990. The association between revisions of financial analysts' earnings forecasts and security-price changes. *Journal of Accounting and Economics* 13: 341–363.

Matsumoto, D. 2002. Management's incentives to guide analysts' forecasts. *The Accounting Review* 77 (3): 483–514.

Loh, R., and G. Mian. 2006. Do accurate earnings forecasts facilitate superior investment recommendations? *Journal of Financial Economics* 80 (2): 455–483.

Mikhail, M., B. Walther, and R. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35: 131–166.

Park, C., and E. Stice. 2000. Analyst forecasting ability and the stock price reaction to forecast revisions. *Review of Accounting Studies* 5 (3): 259–272.

Penman, S. 2001. *Financial Statement Analysis and Security Valuation*. New York, NY: McGraw-Hill/Irwin.

Porter, M. 1980. *Competitive Strategy Techniques for Analyzing Industry and Competitors*. New York, NY: The Free Press.

Rees, L., and S. Sivaramakrishnan. 2007. The effect of meeting or beating revenue forecasts on the association between quarterly returns and earnings forecast errors. *Contemporary Accounting Research* 24 (1): 259–290.

Schipper, K. 1991. Commentary on analysts' forecasts. *Accounting Horizons* 5 (4): 105–121.

Sinha, P., L. Brown, and S. Das. 1997. A re-examination of financial analysts' differential earnings forecast accuracy. *Contemporary Accounting Research* 14: 1–42.

Soffer, L., R. Thiagarajan, and B. Walther. 2000. Earnings preannouncement strategies. *Review of Accounting Studies* 5 (1): 5–26.

Stickel, S. 1990. Predicting individual analyst earnings forecasts. *Journal of Accounting Research* 28 (2): 409–417.

———. 1991. Common stock returns surrounding earnings forecast revisions: More puzzling evidence. *The Accounting Review* 66 (2): 402–416.

———. 1992. Reputation and performance among security analysts. *The Journal of Finance* 47 (6): 1811–1836.

Wasley, C., and J. Wu. 2006. Why do managers voluntarily issue cash flow forecasts? *Journal of Accounting Research* 44 (2): 389–429.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 2075–2105

American Accounting Association
DOI: 10.2308/accr.2010.85.6.2075

Investor Reaction to Going Concern Audit Reports

Krishnagopal Menon
Boston University

David D. Williams
The Ohio State University

ABSTRACT: The literature provides mixed evidence on whether investors find audit reports modified for going concern reasons to be useful. Using a substantially larger sample than previous studies, we observe negative excess returns when the going concern audit report (GCAR) is disclosed. We find that the reaction is more negative if the GCAR cites a problem with obtaining financing, suggesting that the GCAR provides new information to investors. Also, the reaction is more adverse if the GCAR triggers a technical violation of a debt covenant that restricts the firm from getting a GCAR. The evidence suggests that institutional investors drive the reaction to the GCAR, since there is no detectable reaction at low levels of institutional ownership. The market reaction gets more negative as the level of institutional ownership increases, and there is a decline in institutional ownership after the GCAR is issued. We attribute these results to sophisticated investors' awareness of the firm's financing needs and the covenants carried by the firm's debt.

Keywords: *audit reports; going concern; debt covenants; institutional ownership.*

Data Availability: *The data used in this study are publicly available through sources identified in the study.*

I. INTRODUCTION

SAS No. 59 (AICPA 1988) requires auditors to assess a client's going concern status. If, after considering all relevant information, the auditor has substantial doubt about the entity's ability to continue as a going concern, then the auditor is required to modify the audit

We acknowledge helpful comments from Anil Arya, Anne Beatty, Steve Kaplan, Eric Spires, Gary Taylor, Jeff Yu, Teri Ziegler, participants at the Boston University workshop, two anonymous reviewers, and Mark DeFond (*Ad Hoc* Editor). We also thank PwC Foundation for their financial support through the PwCInquires program.

Editor's note: Accepted by Mark DeFond, special editor.

Submitted: January 2008
Accepted: April 2010
Published Online: November 2010

report.[1] We refer to an audit report modified for going concern reasons as a "going concern audit report" (GCAR). There is a long-standing debate on the usefulness of the GCAR. Auditors have expertise in assurance audits, not in judging the going concern status of a firm, and their assessment may not add to what investors already know. On the other hand, auditors have access to information unavailable to investors and can reveal this information through the GCAR.

One way of judging the usefulness of the GCAR is to see whether investors react to GCAR announcements. If the auditor's unfavorable assessment provides new information on the firm, then the price of the firm's securities should decline. Several event studies have examined the market reaction to the issuance of GCARs.[2] While early studies (Elliott 1982; Dodd et al. 1984) do not detect a negative market reaction to qualified reports, including GCARs, later studies (Fleak and Wilson 1994; Jones 1996) find a negative reaction to GCARs. More recent research by Herbohn et al. (2007), using an Australian sample, and Blay and Geiger (2001) do not report a significant negative reaction to GCARs.

Our study provides evidence on this issue, using a substantially larger sample than those used in previous studies. We find, on average, a significant negative reaction to the announcement of a first-time GCAR. However, our primary contribution to the literature lies in going beyond the initial question of *whether* investors react to the GCAR. We investigate *why* investors react to the GCAR and *who* (which class of investors) reacts to the GCAR.

One reason a GCAR may elicit a reaction from investors is that it can provide new information on the status of the client's negotiations with lenders and the client's plans to raise financing. Sometimes, auditors specifically refer to financing problems in the GCAR. We find that investors react more adversely to GCARs that cite financing problems. Further, since many firms carry debt with a covenant that requires the firm to present GCAR-free financial statements, the GCAR can a trigger covenant violation that results in stricter loan terms. We show that the reaction to the GCAR is worse for firms that violate a covenant when they receive a GCAR.

We also consider whether the reaction to the GCAR depends upon the firm's ownership composition. Kausar et al. (2006) present evidence that institutional investors divest shares in firms receiving a GCAR. We find that the price reaction to the GCAR varies inversely with the level of institutional ownership. Also, the negative market reaction for firms with GCARs citing financing problems or for firms violating a debt covenant by receiving a GCAR is observed only at higher levels of institutional ownership. We attribute the differences to institutional investors likely having greater knowledge of the terms of the firm's debt contracts.

Our study contributes primarily to the line of auditing research on GCARs. First, we provide relatively unambiguous evidence that, on average, going concern audit reports provide useful information. Second, we provide insight into the reasons that investors find the GCAR to be useful by showing that GCARs that cite financing problems result in a more negative reaction than do other GCARs. Third, while there is evidence of the importance of restrictive debt covenants in other areas of accounting research, there is little or no evidence in auditing. We show that audit-related covenants also affect investors' decisions. Fourth, our study adds to the literature by showing that institutional investors use the information in audit reports.

In the next section, we discuss the usefulness of GCARs and review the relevant literature. We expand on the reasons why the report should be useful to investors and we develop hypoth-

---

[1] The Financial Accounting Standards Board has an ongoing project intended to make managers, not just auditors, responsible for identifying and disclosing going concern uncertainties.

[2] For convenience, we use the term "GCAR" to refer not only to audit reports modified for going concern purposes, the appropriate terminology under SAS No. 59, but also to "going concern qualifications," the terminology used before SAS No. 59.

eses. In Section III, we present the sample, descriptive statistics, hypothesis tests, and results. Section IV includes additional analyses and robustness tests. In Section V, we present some concluding comments.

II. ARE GOING-CONCERN AUDIT REPORTS USEFUL?

SAS No. 59 (AICPA 1988) directs auditors to actively assess whether the client is a going concern, and, if there is substantial doubt about the client's going concern status, to modify the audit report accordingly.[3] The usefulness of GCARs has long been questioned, even within the accounting profession (Bellovary et al. 2006). Auditors have unique training and expertise in audits, not in assessing the firm's future financial prospects. Further, investors are able to form their own assessments of the likelihood that the firm will be a going concern using a variety of information sources, including financial statements. At the same time, auditors have access to insider information (such as management plans and the status of client-lender negotiations) that could allow them to form superior judgments to those formed by investors who do not have access to similar information.

Several studies have investigated the market reaction to GCARs.[4] Some of these studies do not look at GCARs separately, but along with other audit report modifications or qualifications. Elliott (1982) studies firms receiving "subject-to" audit qualifications and finds no significant market reaction to the release of the audit qualification. Dodd et al. (1984) also find no evidence that investors react adversely to audit qualifications, although a follow-up paper by Dopuch et al. (1986) reports that firms that have qualified opinions that are reported in the business press experience negative excess returns around the report date.

Jones (1996) observes negative returns around the announcement date for 68 firms receiving GCARs, with the return worsening for more unexpected GCARs. Similarly, Fleak and Wilson (1994) find that 114 firms receiving GCARs experience negative returns when compared with distressed firms receiving unmodified opinions. In contrast, Blay and Geiger (2001) find a statistically insignificant market reaction, on average, to a sample of 121 firms receiving GCARs.[5] In a study of 229 Australian firms, Herbohn et al. (2007) find no short-term market reaction to the release of the annual report containing the GCAR.[6]

In summary, results are mixed on the question of whether investors react adversely to the issuance of going concern audit reports. We expect a GCAR to result in a negative investor reaction for two reasons. First, the auditor may provide new information to investors in the GCAR. While investors can extract information about the firm's financial condition efficiently from financial statements, SAS No. 59 directs auditors to go beyond the financial statements in assessing the client's going concern status and to look at management's plans to address liquidity, including

---

[3] Ellingsen et al. (1989) provide a detailed comparison between SAS No. 59 and its predecessor, SAS No. 34, which merely required auditors to ensure that going concern uncertainties that come to light during the normal course of an audit are adequately disclosed.

[4] Reports in the business press commonly claim that the disclosure of a GCAR results in a negative investor reaction (e.g., Carns 2000).

[5] Blay and Geiger (2001) observe a difference in returns between firms in their sample that subsequently went bankrupt and firms that survived. The authors attribute their results to differential market expectations for the two groups; that is, investors may already have anticipated impending bankruptcy for firms that went bankrupt following the GCAR.

[6] Of the event studies cited here that use samples of firms receiving GCARs, Jones (1996), Blay and Geiger (2001), and Herbohn et al. (2007) study only first-time GCARs; Fleak and Wilson (1994) include both first-time and repeat GCARs. Some studies treat the first release of information (either a press release or the filing of the 10-K) on the GCAR as the event date (Blay and Geiger 2001; Herbohn et al. (2007)); others use the annual report or 10-K filing date as the event date. Blay and Geiger (2001) use size-adjusted returns to measure the market's reaction to the event; Herbohn et al. (2007) use market-adjusted returns; other studies use risk-adjusted returns.

financing plans, and the likelihood that these plans will adequately address the firm's problems.[7] Much of this information is not publicly available, and by issuing a GCAR the auditor may reveal this information and change investors' beliefs. Hopwood et al. (1989) show that the audit report provides a statistically significant contribution to a bankruptcy prediction model. Also, Chen and Church (1996) and Holder-Webb and Wilkins (2000) find that GCARs reduce the surprise associated with bankruptcy announcements, suggesting that investors incorporate the information in a GCAR in assessing the likelihood of bankruptcy.

Second, we argue that GCARs can impose contracting costs on firms, which is likely to result in an adverse market reaction. We focus on GCARs that result in technical violation of debt covenants. We discuss this explanation later in this section.

Where most previous studies of GCARs are limited by relatively small sample sizes, our study uses a large sample of 1,194 firms that receive first-time GCARs. Improved data availability allows us to identify the first time that investors learn of the GCAR more precisely than done in many previous studies. With these improvements, we expect to show that investors value the auditor's assessment of the firm's going concern status. Our first hypothesis, in the alternative form, follows.

H1: Investors react negatively to the announcement of a GCAR.

While our paper contributes to the literature in providing large-sample evidence on the GCAR's usefulness to investors, its main contribution lies in exploring reasons for the market's reaction. We investigate whether the informativeness of a GCAR depends upon the reason provided by the auditor for issuing the modified report, the presence of a GCAR-restricting debt covenant, and the level of institutional ownership.

In issuing a GCAR, the auditor reveals his/her assessment of the likelihood that the firm will survive beyond a year, and in doing so may implicitly reveal information about financing plans and the status of the firm's dealings with lenders and others. Sometimes, however, the auditor explicitly provides information on the firm's difficulties in servicing debt or in raising capital. In a GCAR, the auditor typically provides an explanation for why the audit report was modified. Often, the reason provided is very general (for example, that the firm has sustained recurring operating losses) and does not reveal anything that is not already apparent from financial statements. However, some audit reports are more specific and attribute the auditor's decision to issue a going concern modification to particular operating or financing problems. If the auditor cites the firm's financing problems, including difficulties in negotiating with existing lenders or problems in obtaining new financing, then we expect that investors will see the GCAR as conveying the auditor's pessimism about the firm's ability to raise needed capital, and will react more negatively to the GCAR than if the GCAR does not mention financing problems.

H2: Investors react more negatively to GCARs that specifically cite financing problems than to GCARs that do not cite financing problems.

A GCAR potentially affects the firm's contracting costs in several ways. It could cause stock exchanges to question whether the firm should continue to be listed.[8] Lenders may adjust the terms of their loans, credit-rating agencies may adjust their ratings, and suppliers may curtail credit to

[7] Norris (2003) provides an anecdote of an instance where the critical determinant in the auditor's decision to issue a GCAR was the likelihood that the client (Polaroid Corp.) would not obtain refinancing.

[8] For example, New York Stock Exchange rules indicate that receiving a GCAR may provide cause for a firm to be delisted. Both the NASDAQ and the American Stock Exchange have recently (since 2004) required listed firms to inform investors of the receipt of a GCAR by issuing a press release within seven days following the 10-K filing.

firms receiving GCARs.[9] In all, the GCAR can severely constrain the firm's ability to raise capital and continue to operate. Investors are likely to anticipate the increased contracting costs that follow GCARs and to react negatively to the announcement of a GCAR.

We focus on the existence of a restrictive debt covenant requiring the firm to prepare financial statements without a GCAR. If the debt carries such a covenant, then the issuance of a GCAR to the borrower would likely lead to higher contracting costs. Beneish and Press (1993) find that technical violation of accounting-related covenants results in higher interest costs, debt-restructuring costs, and increased lender control. As a result, investors react negatively to reports of technical violations of covenants (Beneish and Press 1995). Since investors are aware of covenants, we expect that violation of a covenant restricting the firm from receiving a GCAR will have an adverse result on stock prices. This leads to our third hypothesis:

H3: Investors react more negatively to GCARs issued to firms that have a debt covenant that restricts the firm from receiving a GCAR than to GCARs issued to firms that do not have such a covenant.

Kausar et al. (2009) find that U.S. firms receiving GCARs experience negative returns in the year following the GCAR disclosures.[10] They attribute this apparent under-reaction on the part of investors to market frictions, specifically high transaction costs associated with trading in GCAR-receiving firms. In a related study, Kausar et al. (2006) find that institutional investors react differently from individual investors to the GCAR. Institutional investors reduce their holdings in firms receiving GCARs.

Prior studies have identified differences in the way that institutional investors and individual investors react to accounting information. Several potential reasons have been identified for this differential reaction. Two common reasons offered are: (1) institutional investors possess information advantages, such as greater predisclosure information (e.g., Utama and Cready 1997); and (2) there are differences in the level of sophistication of institutional and individual investors, stemming in part from differences in the amount of time and resources that these two investor groups can afford to spend on their investment decisions (e.g., Walther 1997).

If institutional investors have an informational advantage with respect to the GCAR, then this advantage should be reflected in these investors knowing about the GCAR before it is announced and reacting before the GCAR is announced, rather than reacting after it is announced. We should then observe the adverse reaction to the announcement of a GCAR being mitigated by increasing levels of institutional ownership. However, even if institutional investors had predisclosure information about the GCAR pre-Regulation FD, it is likely that this advantage would be largely dissipated following Regulation FD. Since most of our sample is drawn from the post-Regulation FD period, this advantage, if it exists, should be muted for our sample.

Even absent information advantages, institutional investors could use their superior resources and expertise to make different investment decisions from individual investors. For example, institutional investors may be better aware of the status and terms of the firm's debt contracts, including whether the debt carries a covenant restricting a GCAR and the likely costs of violating this covenant.

---

[9] For example, in 2002, Standard & Poor's cut GenTek Inc.'s corporate credit rating, citing the company's poor financial performance and receipt of a GCAR (*Business Recorder* 2002). Landler (2004) cites a company (Spiegel Inc.) that delayed filing its 10-K because disclosing that the firm had received a GCAR "would have jeopardized Spiegel by unnerving its suppliers and disrupting [the CEO's] negotiations with the banks."

[10] Similarly, Taffler et al. (2004) compute buy-and-hold abnormal returns for a portfolio of U.K. firms in the 12 months following their receipt of a GCAR, and find that these firms substantially underperform reference portfolios. Ogneva and Subramanyam (2007), in contrast, using a sample of U.S. firms, suggest that the Taffler et al. (2004) results depend on the model used to generate abnormal returns.

Note that the information advantage and investor sophistication arguments point to different predictions for the relationship between the level of institutional holdings and the announcement period reaction to the GCAR. If the investor sophistication argument prevails, then we expect institutional investors to react more negatively to the GCAR than individual investors, and the market reaction to the GCAR should become more adverse with increasing levels of institutional holdings. We believe the investor sophistication argument is more persuasive and we express H4 accordingly.

H4: The investor reaction to the announcement of a GCAR is negatively related to the level of institutional ownership in the firm.

III. METHOD AND TESTS

The sample for the study consists of firms that receive an initial GCAR in the period 1995 to 2006. We identify firms receiving an initial GCAR (that is, firms receiving a GCAR that did not receive a GCAR in the previous year) by searching the Compact Disclosure database, the Audit Analytics database, Compustat, and the business media, and then confirming the existence of a first-time GCAR by reading audit reports in current and previous 10-K filings. We exclude all American Depositary Receipts (ADRs). We include only the first GCAR received by the firm in our sample, even if the firm obtains a GCAR, or several consecutive GCARs, is then issued an unmodified opinion, and later receives a GCAR once more. Consequently, each firm appears only once in our sample. We apply this rule since we believe market expectations and security price reactions will be different for firms receiving continuing GCARs.

Our initial sample consists of 4,541 firms with first-time GCARs. Table 1, Panel A details how we arrived at the final sample. We exclude firms that do not have data on CRSP and Compustat. We also omit firms that are already in bankruptcy. We drop very small firms (firms with less than $1 million in either total assets or market value of equity).[11] Our final sample consists of 1,194 firms. Table 1, Panel B provides the frequency of observations by year. The highest frequency of GCARs is observed in the year 2000.

Table 2 provides descriptive statistics for the firms in the sample. Market prices are obtained from CRSP, institutional ownership data are obtained from the CDA/Spectrum Institutional Holdings database, and filing dates are obtained through the SEC's EDGAR database or Lexis-Nexis. On average, sample firms receiving GCARs are small (mean total assets = $165.7 million; mean market value of equity = $50.7 million). The firms are very distressed, as should be expected for GCAR-receiving firms, with negative and declining profitability, negative cash generating ability, and high leverage. The level of institutional ownership averages 13.0 percent. SEC rules specify that firms must file their 10-K reports within a certain number of days of the end of the fiscal year. We find that nearly one half (48.9 percent) of the sample have material delays in filing their 10-K, which we define as filing more than five days late; 68.2 percent of the sample firms have Big 4 auditors. We also calculate the earnings surprise for the 497 firms in the sample for which we have analyst forecast data. The earnings surprise, scaled by stock price, has a mean of −0.62 and a median of −0.07. The low mean likely reflects the nature of the GCAR sample (highly distressed firms with low stock prices).

Market Reaction to First GCAR Disclosure

We use the earliest date on which the GCAR is announced to investors as the event date to determine whether there was a market response to the GCAR announcement. Some firms an-

[11] The excluded small firms have the following characteristics (all medians): total assets: $1.52 million; market value of equity: $0.58 million; earnings before interest and taxes: $(2.60) million; stock price: $0.21 per share.

TABLE 1

Sample Selection and Distribution by Year

Panel A: Sample Selection

Sample	Frequency
Firms receiving initial going concern audit reports from 1995–2006	4,541
− Firms without required stock price data	2,130
− Firms without required financial data	1,044
− Firms with assets or market value under $1 million	96
− Firms that declare bankruptcy prior to the filing of Form 10-K	77
Full Sample	1,194

Panel B: Sample Distribution by Year

Year	Frequency	Percentage
1995	66	5.5%
1996	109	9.1%
1997	101	8.5%
1998	131	11.0%
1999	79	6.6%
2000	169	14.2%
2001	133	11.1%
2002	121	10.1%
2003	70	5.9%
2004	67	5.6%
2005	77	6.4%
2006	71	6.0%
Total	1,194	100.0%

nounce the issuance of the GCAR prior to the release of the 10-K through a press release ("Early Disclosers"), but in most cases investors first learn of the GCAR in the 10-K ("10-K Disclosers"). For Early Disclosers, we use the press release date as the event date.[12] For 10-K Disclosers, we use the 10-K filing date as the event date, since this is the first date that investors become aware of the GCAR, absent information leakage.

To identify firms that belong to the Early Disclosers sample, we review (through Lexis-Nexis) all press releases placed on newswires by each GCAR firm prior to the filing of the 10-K announcing the GCAR. We find that 105 of the 1,194 firms issue a press release at least three days

[12] It is unclear why some firms make early disclosures of GCARs. Early disclosers in our sample are not different from other sample firms in terms of profitability, financial distress, or in the likelihood they employ a Big 4 auditor. However, they tend to be larger than the other firms and to have more analysts following them. Early disclosers also are more likely to have made an accounting restatement in the past year. The larger size and analyst following may be consistent with arguments in the voluntary disclosure literature that managers voluntarily disclose bad news because of reputation and litigation concerns (e.g., Skinner 1994).

TABLE 2

Descriptive Statistics

Full Sample (n = 1,194)	Mean	Standard Deviation	Lower Quartile	Median	Upper Quartile
Assets ($ millions)	165.7	758.9	9.0	25.5	90.9
Market value of equity ($ millions)	50.7	100.5	7.5	19.0	47.4
Earnings before interest and taxes	−0.60	0.91	−0.81	−0.29	−0.06
Change in income	−0.27	0.79	−0.44	−0.16	0.00
Cash flow from operations ($ millions)	−7.15	48.9	−10.45	−3.22	−0.20
Return on Assets	−0.83	1.20	−1.06	−0.49	−0.19
Z-score	2.03	3.27	0.54	1.27	3.40
Exit Value ($ millions)	−2.75	90.4	−1.71	2.32	11.59
Earnings Surprise (n = 497)	−0.62	4.16	−0.47	−0.07	0.01
Debt/Assets	0.77	0.76	0.45	0.70	0.93
Institutional Ownership	0.13	0.16	0.01	0.07	0.20
Late 10-K filers	48.9%	—	—	—	—
Big 4 auditors	68.2%	—	—	—	—

Total assets, market value of equity, earnings before interest and taxes, change in income, cash flow from operations, return on assets, Exit Value, and debt/assets are all based on year-end or annual numbers for the period for which the GCAR is issued. Change in income is scaled by total assets. Earnings before interest and taxes is scaled by assets and adjusted for the industry average. Return on assets is net income/total assets. Z-score is a composite measure of financial distress based upon Zmijewski (1984). Exit Value is the firm's anticipated liquidation value as computed based upon Berger et al. (1996). Earning Surprise is computed as (actual EPS − mean forecasted EPS)/stock price at beginning of the fourth quarter. Debt/Assets is total liabilities divided by total assets. Institutional ownership is the fraction of outstanding common stock owned by institutions. Late 10-K filers indicates the percentage of firms that file 10-Ks at least five days after the statutory filing date. Big 4 auditors indicates if the firm is audited by a Big 4 audit firm or one of its predecessors.

prior to the 10-K date, indicating that their auditors will issue a GCAR for the current year. Many of these disclosures are provided in early earnings announcements.[13] On average, these firms disclose the GCAR in a press release 11.6 days before their 10-K filing.

In Table 3, we report daily size-adjusted excess returns for the event date and the five days preceding and succeeding the event date for the 1,194 sample firms.[14] Mean excess returns for Day 0, Day +1 and Day +2 are significantly negative, with the most negative return taking place on Day +1 (−3.49 percent).[15] The cumulative three-day excess return (0, +2) is −6.28 percent.[16]

[13] A very small number of firms issue press releases that announce only the GCAR. These firms provide a "pure" sample, in the sense that there is no other information included in the press release that may explain investors' reaction. However, there are too few of these firms to provide any detailed analysis. In the full sample, 35 firms issue press releases containing GCAR announcements but not earnings news. Only 16 of these firms issue the press release either before or concurrent with (up to one day following) the 10-K date. The mean three-day cumulative excess return (0, +2) for the 16 firms issuing early or timely press releases is significantly negative (−8.52 percent, $p < 0.05$), but the mean cumulative excess return for the firms issuing late press releases is not significantly different from zero.

[14] Daily excess returns are computed by taking the firm's daily return and subtracting the corresponding size-decile portfolio return for the day from CRSP. We choose to use size-adjusted returns because the sample is dominated by small firms. A long line of studies, stretching back to Banz (1981), shows a relationship between firm size and market returns. Standard Student's t-statistics are used to test whether the computed daily excess return are significantly different from zero.

[15] It is possible that the larger magnitude of negative returns obtained on day $t+1$, relative to day t, is explained by the availability of the information to investors. For the firms in our sample for which the event date is the 10-K filing date, the information may have been available to investors only the next day. Similarly, some firms announcing GCARs in press releases may have made the disclosure after markets closed for the day.

TABLE 3

Daily Excess Returns Related to the GCAR Disclosure

	All Firms (n = 1,194)				10-K Disclosers Sample (n = 1,089)			
Day	Mean	Std. Dev.	Median	% Negative	Mean	Std. Dev.	Median	% Negative
−5	0.0059	0.1082	−0.0019	51.4%	0.0066	0.1098	−0.0022	52.2%
−4	0.0023	0.1029	−0.0004	51.2%	0.0029	0.1033	−0.0003	50.8%
−3	−0.0017	0.1064	−0.0011	52.1%	−0.0015	0.1071	−0.0018	51.7%
−2	0.0011	0.1324	−0.0021	52.9%	0.0009	0.1355	−0.0023	52.8%
−1	0.0004	0.1164	−0.0020	52.5%	0.0008	0.1139	−0.0017	52.1%
0	−0.0150**	0.1301	−0.0080	59.2%**	−0.0113**	0.1289	−0.0076	58.4%*
1	−0.0349**	0.1353	−0.0179	63.0%**	−0.0323**	0.1315	−0.0174	63.2%**
2	−0.0129**	0.1217	−0.0074	58.4%**	−0.0112**	0.1202	−0.0073	57.9%*
3	0.0042	0.1196	−0.0017	52.9%	0.0038	0.1128	−0.0010	51.7%
4	0.0049	0.1296	−0.0010	51.5%	0.0039	0.1304	−0.0014	52.9%
5	0.0004	0.1135	−0.0035	53.0%	0.0002	0.1084	−0.0029	54.2%

	Early Disclosers Sample (n = 105)			
Day	Mean	Std. Dev.	Median	%Negative
−5	−0.0009	0.0901	0.0007	47.6%
−4	−0.0041	0.0981	−0.0007	52.3%
−3	−0.0043	0.1002	0.0031	49.5%
−2	0.0031	0.0953	−0.0013	54.2%
−1	−0.0034	0.1409	−0.0084	57.1%
0	−0.0543**	0.1367	−0.0199	67.6%*
1	−0.0622**	0.1676	−0.0298	61.0%
2	−0.0299*	0.1355	−0.0091	63.3%
3	0.0084	0.1742	−0.0077	57.1%
4	0.0152	0.1217	0.0026	45.7%
5	0.0030	0.1577	−0.0086	54.3%

*, ** $p < 0.05$, and $p < 0.01$, respectively, two-tailed.

The event date is the earliest GCAR disclosure date (for 10-K Disclosers, the 10-K filing date; for Early Disclosers, the GCAR-disclosing press release date). The table presents statistics for daily size-adjusted excess returns. Significance levels are for t-tests for differences in the mean daily excess return from zero or for Chi-square tests for differences in the percentage of negative excess returns in a given day from the average for the 21 days around the event date for sample firms.

The average percentage of firms incurring negative excess returns each day over the 21 days from Day −10 through Day +10 (where Day 0 is the event date) is 54.0 percent. On Day 0, Day +1, and Day +2, the percentages of firms with negative excess returns are 59.2 percent, 63.0 percent, and 58.4 percent, respectively, all significantly higher than the 21-day average. Table 3 reports daily returns separately for the samples of 10-K Disclosers and Early Disclosers, as well. The

[16] The cumulative excess return for the announcement period (0, +2) is computed by adding the daily excess returns for the three days beginning with the event day.

10-K Disclosers have a cumulative three-day excess return of −5.48 percent.[17] The Early Disclosers experience a mean cumulative three-day (0, +2) excess return of −14.64 percent.

The results shown in Table 3 suggest a negative investor response to the GCAR announcement, consistent with H1.[18] However, the computed announcement period return may be contaminated by other information that is released at the same time. Like other studies investigating the market reaction to GCARs, we are unable to precisely separate the GCAR news from concurrent disclosures. For the majority of firms in the sample, the GCAR news becomes available simultaneously with other 10-K information. Even for firms that announce the GCAR early in a press release, there is generally other information released at the same time, such as annual earnings. In later analyses, we attempt to mitigate this problem and provide greater confidence that the observed negative reaction is due to the GCAR.

Cited Reasons for the GCAR

The GCAR generally carries a brief explanation for why the auditor chose to modify the audit report. We review the explanations for the GCARs in our sample and classify explanations into four categories: (1) *Poor financial performance*. This category includes explanations that cite the financial condition or performance of the company, including references to the firm's present or recurring losses, negative cash flows, poor working capital situation, or negative equity. (2) *Financing problems*. In this category, we include references to debt defaults, debt covenant violations, loss of credit facilities, or the need for, or problems in securing, additional financing.[19] (3) *Operating problems*. This category includes references to specific operating problems, such as the loss of a key customer, problems with suppliers, or business combination issues.[20] (4) *Other*. All other reasons cited for the GCAR are included in this category (such as litigation and regulatory issues).

Table 4, Panel A shows the frequency with which the auditor attributes the GCAR issuance to each of these reasons. Since a GCAR can cite more than one reason, the total frequency adds up to more than 1,194 for the full sample. Only in 15 cases does the auditor not provide a reason for the GCAR. The most commonly cited reason is poor financial performance (75 percent of firms), but in a significant number of cases the auditor cites financing problems (35.5 percent) or operating problems (21.2 percent). Table 4 also breaks out the frequency separately for the 10-K Disclosers sample and the Early Disclosers sample. The most notable difference between these two samples is the higher frequency of early disclosing firms, where the auditor cites financing problems (52.4 percent, against 33.9 percent for the 10-K Disclosers sample).

These different reasons for the GCAR cited by the auditor likely have different information content for investors. A citation of poor financial performance, for example, does not go beyond

[17] Twelve firms in the sample announce the GCAR in press releases on the day preceding or two days preceding the 10-K filing. We include these firms in the 10-K Disclosers sample since the event window includes the 10-K filing. Excluding these firms does not alter any of the reported results for the 10-K Disclosers sample.

[18] The investor reaction to the GCAR announcement is conditioned upon the extent to which the announcement was anticipated. Investors may anticipate the GCAR because of the information they have from interim financial reports and other sources about the financial condition of the firm. To the extent that the GCAR is anticipated, it will bias against observing a negative reaction, since investors will already have priced in the information.

[19] As an example of a financing problem, the auditor's report for Blockbuster Inc. (not in the sample) for the fiscal year 2008 states the following: "[T]he risk the Company may not successfully complete a refinancing of its credit facility scheduled to mature in August 2009 and obtain related amendments of financial covenants included therein, and/or the risk the Company may not have adequate liquidity to fund their operations raise substantial doubt about the Company's ability to continue as a going concern."

[20] An example of a GCAR attributed to an operating problem is provided by Pinnacle Airlines Corp.'s GCAR for the fiscal year 2005. The auditor notes: "[T]he uncertainty regarding the future of the Company's contract with its primary customer due to this customer seeking protection under Chapter 11 of the U.S. Bankruptcy Code raises substantial doubt about the Company's ability to continue as a going concern."

TABLE 4

Reasons Cited for GCARs and Debt Covenant Requirements Related to GCARs

Panel A: Reasons Cited for GCAR in Audit Report/Press Release

	All Firms (n = 1,194)		10-K Disclosers Sample (n = 1,089)		Early Disclosers Sample (n = 105)	
	Frequency	%	Frequency	%	Frequency	%
Poor financial performance	895	75.0	817	75.0	78	74.3
Financing problems	424	35.5	369	33.9	55	52.4
Operating problems	253	21.2	233	21.4	20	19.0
Other	131	11.0	121	11.1	10	9.5
No reason cited	15	1.3	12	1.1	3	2.9

Panel B: Analysis of Debt Covenant Requirements

	All Firms (n = 1,194)		10-K Disclosers Sample (n = 1,089)		Early Disclosers Sample (n = 105)	
	Frequency	%	Frequency	%	Frequency	%
Covenants requiring audit report without GCAR	193	16.2	157	14.4	36	34.3
Covenants requiring "unqualified" report	92	7.7	85	7.8	7	6.7
Covenants with financial statements required to be presented on a going concern basis	21	1.8	19	1.7	2	1.9
Covenants with other restrictions related to financial statements ("presents fairly," "fair and accurate," "in good faith")	49	4.1	44	4.0	5	4.8
Covenants requiring audited financial statements without additional restrictions on auditor's report or financial statements	418	35.0	392	36.0	26	24.8
Covenants requiring audited financial statements	773	64.7	697	64.0	76	72.4
Firms without covenants requiring financial statements	421	35.3	392	36.0	29	27.6
Total	1,194	100.0	1,089	100.0	105	100.0

(continued on next page)

Panel A shows the frequency with which the auditor cites certain reasons for issuing the GCAR in the audit report (for the 10-K Disclosers Sample) or the press release (for the Early Disclosers Sample). The percentage is based on the sample size. Since multiple reasons can be cited in a GCAR, the aggregate of the frequencies exceeds the sample size. Panel B shows the frequency with which debt covenants require firms to prepare audited financial statements as well as the frequencies of restrictions on the audit report.

what can be ascertained from the financial statements and is unlikely to convey new information to investors. In contrast, if the auditor indicates that the firm is having problems related to its financing, the disclosure may be incrementally informative to investors. Operating problems and other specific reasons cited by auditors may also be news to some investors, but in general we expect that these will already be known to the market. We focus on financing problems in our analysis.

Debt Covenants and GCARs

In order to study the contracting consequences of the GCAR, we focus on debt covenants. We search the Dealscan database and the material contracts exhibit in 10-K filings, using Lexis/Nexis to identify covenants related to the GCAR. Table 4, Panel B reports the results. We are unable to find any restrictive covenants related to financial statements or the auditor's report for 421 firms in the full sample.[21] In the remaining 773 instances, the firm is required to submit timely audited financial statements to creditors.

Of the firms that are required to provide audited financial statements, 193 firms are subject to a requirement that the audit report should not be qualified or modified for going concern reasons. In 92 additional cases, the requirement is for an "unqualified" audit report. Since a GCAR is an unqualified audit report according to auditing standards, we do not view these 92 covenants as restricting GCARs. Covenants in the remaining 488 firms require timely audited financial statements, but impose no restrictions on the audit report. Some of these firms have special wording on the financial statements: 21 firms require that the financial statements be prepared on a going concern basis; 49 firms have other wording with respect to the quality of financial statements (e.g., "presents fairly," "fair and accurate," or "in good faith").

In our analysis, we classify only the 193 firms that are required to present an audit report that is not modified for going concern reasons as having a restrictive covenant prohibiting the GCAR. Descriptive statistics (untabulated) show that these firms tend to be larger (mean assets = \$272.3 million) than the other firms in the sample (mean assets = \$148.6 million) and to have lower levels of losses (mean return on assets = −0.664, as compared to −0.857). In addition, they are much more likely to have Big 4 auditors (82.3 percent versus 65.4 percent).

Table 4, Panel B also breaks down the debt covenant requirement separately for the 10-K Disclosers and the Early Disclosers samples. A higher percentage of the early disclosing firms (72.4 percent versus 64.0 percent) have covenants that require the firm to prepare audited financial statements, and a substantially higher percentage (34.3 percent against 14.4 percent) have covenants that restrict the firm from getting a GCAR.

Multivariate Analyses

We use OLS regressions to test our hypotheses related to the informativeness of GCARs. We begin by controlling for several additional variables that could explain the investor reaction on the event date. We estimate the following model:

$$CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \varepsilon. \quad (1)$$

The event date is the GCAR disclosure date. The dependent variable (*CAR*) measures the cumu-

[21] Using Lexis-Nexis, we search Exhibit 10 (material contracts) of 10-K filings for all companies in the sample starting in 1993. We expect to detect covenants in any debt taken in 1993 or later. However, it is possible that some firms in our sample had debt prior to 1993 that included covenants related to financial statements, and that we fail to detect these firms. These firms are misclassified in the analysis, and the potential misclassification biases against our finding significant results in tests for *GCCovenant*.

lative size-adjusted excess return over the three days beginning with the event date, which is the date the GCAR is first announced. The independent variables *LMValue*, *EBIT*, *δIncome*, *CFOPS*, *Zscore*, and *ExitValue* capture aspects of financial performance or condition that may explain the market's reaction to GCAR-accompanying disclosures. *LateFiler* is an additional control variable that may explain the market's reaction in the event window. Consistent with H1, we expect a negative intercept, which would indicate that the market reacts negatively to the GCAR even after controlling for contemporaneous disclosures of financial condition.

LMValue, the natural log of the market value of equity, is a proxy for firm size, and is measured at the end of the fiscal year for which the firm receives a GCAR, that is, several weeks before the GCAR is announced. This variable controls for size effects not fully captured by our use of size-adjusted returns. We expect firm size to proxy for the firm's information environment. Since larger firms are likely to have richer information environments, we would expect less surprise associated with their GCAR disclosures, resulting in a positive coefficient for *LMValue*. *EBIT* is the firm's earnings before interest and taxes, scaled by total assets, less the industry mean; *δIncome* is the change in net income scaled by total assets; *CFOPS* is cash flow from operations scaled by total assets. *EBIT* and *CFOPS* are measured using annual numbers for the period for which the GCAR is issued, while *δIncome* is the change in net income from the previous year to the GCAR year.[22] *Zscore* is a measure of financial distress, computed using the Zmijewski (1984) distress score metric.[23] In general, we expect the dependent variable to be positively associated with financial performance, so coefficients for *EBIT*, *δIncome*, and *CFOPS* should be positive and the coefficient for *Zscore* should be negative.

ExitValue represents the exit value of assets, computed according to the Berger et al. (1996) formula, scaled by the market value of equity. *ExitValue* is measured at the end of the fiscal year for which the GCAR was issued.[24] We expect investors to revise the probability of bankruptcy upward in highly distressed companies and consequently to place greater weight on the exit value (e.g., Barth et al. 1998; Collins et al. 1999). As a result, β_6 should be positive.

LateFiler is a dichotomous variable identifying firms that file the 10-K after the SEC-specified filing date. *LateFiler* is assigned a value of 1 if the firm's 10-K filing is more than five days late.[25] Late filing companies are known to be poorer performers (e.g., Alford et al. 1994). If the firm files a late 10-K, then the market will likely anticipate that the financial statements contain bad news, and perhaps even that the audit process has been extended because a GCAR is being contemplated. As a result, some of the bad news would already have been impounded in the stock price during the period of the delay, resulting in a less adverse reaction to the 10-K filing. We would then expect β_7 to be positive. We drop *LateFiler* in estimating the model for the Early Disclosers sample since the late filing has not yet happened at the event date for this sample.

We add five variables to Model (1) in our second specification. Three of these variables (*CitedDebt*, *GCCovenant*, and *InstOwn*) allow us to test H2, H3, and H4, respectively.

---

[22] Some year-end values may not be available to investors on the event date for the Early Disclosers sample. Still, we use year-end values for the Early Disclosers firms because a substantial majority of the press releases disclosing the GCAR early also provide annual earnings and balance sheet data.

[23] The Zmijewski (1984) model includes three variables: return on assets, debt/assets, and the current ratio.

[24] Berger et al. (1996) derive the following formula: *Exit Value* = 1.0 * *Cash* + 1.0 * *Marketable Securities* + 0.72 * *Receivables* + 0.55 * *Inventory* + 0.54 * *Fixed Assets* − 1.0 * *Payables* − 1.0 * *Total Debt*. Like Berger et al. (1996) and Collins et al. (1999), we replace negative estimated values with zero, since a negative exit value has no economic meaning for a limited liability company.

[25] Until 2002, all firms were expected to file their 10-K reports within 90 days from the fiscal year-end. Starting from December 15, 2002, this period was shortened to five days for accelerated filers (mainly firms with common equity float exceeding $75 million). In 2006, the filing period was shortened further for large firms.

$$CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \beta_8 Big4 + \beta_9 CitedDebt + \beta_{10} FSCovenant + \beta_{11} GCCovenant + \beta_{12} InstOwn + \varepsilon. \quad (2)$$

We include *Big4* (1 if the firm employs a Big 4 auditor, and 0 otherwise) because prior studies have shown that investors view Big 4 auditors as being more credible than smaller auditors.[26] For example, Geiger and Rama (2006) find that Big 4-issued audit reports are more accurate in bankruptcy prediction (that is, there are fewer Type I and Type II errors) than reports issued by smaller auditing firms. It is plausible that investors' reactions to the GCAR depend upon the credibility of the issuing auditor. Thus, investors may see it as more likely that a firm will go bankrupt if it receives a GCAR issued by a Big 4 auditor, leading to a more adverse market reaction for Big 4-issued GCARs.

CitedDebt represents the reason for the GCAR cited in the audit report. *CitedDebt* equals 1 if the auditor attributes the GCAR to a financing problem and 0 otherwise. A negative coefficient will support H2, indicating that investors react more adversely to GCARs issued because of financing problems.

FSCovenant and *GCCovenant* represent financial statement-related covenants. *FSCovenant* equals 1 if the firm has lending agreements that require timely audited financial statements (and 0 otherwise). *GCCovenant* is coded as 1 if a covenant specifies that the firm should not receive a GCAR (and 0 otherwise).[27] Although we do not have a hypothesis related to *FSCovenant*, we expect the coefficient for *FSCovenant* to be negative. Lenders requiring audited financial statements are likely to be stricter than lenders that do not require audited statements. A negative coefficient for *GCCovenant* would indicate that firms for which the GCAR violates a covenant experience an even more negative investor reaction than firms whose lenders just require financial statements, supporting H3.

InstOwn indicates the percentage of outstanding common stock held by institutions in aggregate. *InstOwn* indicates the level of institutional holdings at the calendar quarter-end preceding the calendar quarter in which the GCAR is announced. Following H4, we expect institutional investors to react more strongly to a GCAR than individual investors. The coefficient for *InstOwn* should be negative. A positive coefficient would be consistent with the information advantage argument dominating the investor sophistication argument.

We estimate Models (1) and (2) on the Early Disclosers sample (105 firms) and the 10-K Disclosers sample (1,089 firms).[28,29] Table 5 reports the regression results. We first discuss the results for the 10-K Disclosers sample. The intercept in Model (1), which includes only variables

[26] We refer to the tier of largest auditors as the Big 4 in this study. The tier includes six firms in the initial years of the sample until the 1998 merger between Price Waterhouse and Coopers & Lybrand, and then five firms until the breakup of Arthur Andersen in 2002.

[27] Note that only firms that are required to present audited financial statements to lenders can have the additional covenant prohibiting a GCAR. As Table 4, Panel B shows, there are 773 firms that are required by lenders to prepare audited financial statements (i.e., *FSCovenant* = 1). Of these, 193 firms are required to have audit reports without a GCAR (i.e., *GCCovenant* = 1), whereas 580 firms are not (i.e., *GCCovenant* = 0).

[28] We winsorize all continuous independent variables used in this regression (and in the other regressions used in this paper) at the 1st and 99th percentiles. The results reported in the tables reflect this winsorizing procedure. In robustness tests, we use alternate winsorizing rules (0.5th and 99.5th percentiles; 2nd and 98th percentiles) and obtain similar results. We also identify observations that have a large effect on regression coefficient estimates (following Belsley et al. 1980) and delete the observations (1 percent of the sample) with the largest effect. The only notable difference in the results is a decline in the significance level for *FSCovenant* ($p = 0.06$).

[29] We test for multicollinearity in all the regressions we estimate in this study, using variance inflation factors. Variance inflation factors are well within acceptable levels, the highest value being 2.96.

TABLE 5

OLS Regressions on the Market Reaction to the GCAR Disclosure

Model 1: $CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \varepsilon$

Model 2: $CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \beta_8 Big4 + \beta_9 CitedDebt + \beta_{10} FSCovenant + \beta_{11} GCCovenant + \beta_{12} InstOwn + \varepsilon$

		10-K Disclosers Sample (n = 1,089)				Early Disclosers Sample (n = 105)			
		Model (1)		Model (2)		Model (1)		Model (2)	
Variable	Exp. Sign	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept		−0.0483	−2.94**	−0.0119	−0.65	−0.2237	−3.91**	−0.1146	−1.73
LMValue	+	−0.0042	−0.96	0.0061	1.33	0.0311	2.15*	0.0455	3.03**
EBIT	+	−0.0134	−1.11	−0.0038	−0.32	0.0312	0.66	0.0215	0.48
δIncome	+	0.0088	1.20	0.0084	1.16	0.0868	3.04**	0.0759	2.74**
CFOPS	+	−0.0086	−0.59	−0.0018	−0.12	−0.0981	−1.54	−0.0467	−0.74
Zscore	−	−0.0027	−1.08	−0.0005	−0.21	−0.0047	−0.55	0.0001	0.01
ExitValue	+	0.0067	0.35	0.0068	0.36	0.1122	1.86	0.1108	1.94
LateFiler	+	0.0174	1.59	0.0141	1.31	—	—	—	—
Big4	−	—	—	−0.0165	−1.39	—	—	0.0005	0.01
CitedDebt	−	—	—	−0.0254	−2.26*	—	—	−0.0774	−1.99*
FSCovenant	−	—	—	−0.0255	−2.24*	—	—	−0.1155	−2.65**
GCCovenant	−	—	—	−0.0434	−2.59**	—	—	−0.1004	−2.10*
InstOwn	−	—	—	−0.1591	−4.31**	—	—	−0.2493	−2.20*
R^2 (adj)		0.007		0.048		0.070		0.123	

(continued on next page)

TABLE 5 (continued)

*, ** $p < 0.05$, and $p < 0.01$, respectively. Significance levels are for two-tailed t-tests.

The dependent variable is the size-adjusted three-day (0, +2) cumulative excess return related to the GCAR disclosure date (for 10-K Disclosers, the 10-K filing date; for Early Disclosers, the GCAR-disclosing press release date).

Variable Definitions:

LMValue = natural log of market value of equity in millions;
EBIT = industry-adjusted earnings before income taxes and interest, scaled by assets;
δIncome = change in net income, scaled by assets;
CFOPS = cash flow from operations/total assets;
Zscore = a composite measure of financial distress based upon Zmijewski (1984);
ExitValue = the firm's exit value, computed based upon Berger et al. (1996);
LateFiler = 1 if the firm's 10-K was filed five days or more after the SEC statutory filing date, and 0 otherwise;
Big4 = 1 if audited by a Big 4 audit firm or one of its predecessors, and 0 otherwise;
CitedDebt = 1 if the GCAR cited a financing issue, and 0 otherwise for the 10-K sample, and 1 if the press release attributed the GCAR to a financing issue, and 0 otherwise for the Early Disclosers sample;
FSCovenant = 1 if the firm has a debt covenant that requires the client to present audited financial statements, and 0 otherwise;
GCCovenant = 1 if the firm has a debt covenant that requires the client to present an auditor's report without an explanatory paragraph noting a going concern limitation, and 0 otherwise; and
InstOwn = the fraction of outstanding common stock owned by institutional investors.

that control for financial information that may explain the market reaction on the event date, is significantly negative, as expected under H1. None of the control variables are significant.

The intercept is not significant with the inclusion of the additional variables in Model (2). The control variables again are not significant, as in Model (1). *CitedDebt* has a negative and significant coefficient, suggesting that GCARs that cite financing problems induce a more negative reaction. This finding supports H2. *FSCovenant* has a significantly negative coefficient, indicating that firms with covenants requiring audited financial statements experience a more negative reaction do firms without such a requirement, but, consistent with H3, the sign and significance of the *GCCovenant* coefficient point to an even more negative reaction if the firm's audit report is required to be free of a GCAR. Finally, consistent with H4, the coefficient for *InstOwn* is negative and significant, showing that the abnormal stock returns in the event window are more negative in firms with higher levels of institutional holdings.

We estimate Models (1) and (2) on the sample of 105 Early Disclosers firms, as well. We make two changes to the models for this sample. *LateFiler* is not included, since we measure returns around the press release date and not the 10-K date. *CitedDebt* is defined differently for this sample, since the actual GCAR is not visible to investors at the time of the press release. For early disclosers, we define *CitedDebt* as 1 (and 0 otherwise) if managers attribute the GCAR to a financing problem when they disclose the impending GCAR in the press release.

The regression results show small differences from the results for the larger sample of firms that disclosed the GCAR only in the 10-K. *LMValue* (firm size) and *δIncome* (the change in profitability) have positive and significant coefficients. Among the variables of interest, *GCCovenant* has a significant coefficient, even in the presence of a significant coefficient for *FSCovenant*, and *CitedDebt* and *InstOwn* are also significant in the expected direction.

Overall, the analysis of returns on the first announcement of the GCAR supports the position that the GCAR is informative to investors. We observe negative returns on the announcement, and these returns continue to be negative after controlling for the firm's financial performance (H1). Returns are more negative for firms where the auditor cites financing problems as being the reason for the GCAR (H2) and for firms that violate a debt covenant by receiving a GCAR (H3). Also, returns are increasingly negative as the level of institutional holdings increases (H4).

Controlling for Earnings Surprise

If firms make an early disclosure of the GCAR and the disclosure is included in an earnings announcement, then investors will likely react to the earnings surprise in the announcement window. If firms do not make an early earnings announcement, then investors may react to the earnings surprise in the 10-K event window. It is important, then, to control for earnings surprise in our analysis. The availability of earnings forecast data proves to be a major constraint. Of the 1,194 firms in our sample, we are able to obtain earnings forecast data from I/B/E/S for only 497 firms. We re-estimate Models (1) and (2) on firms with earnings forecast data ("partial sample") to see whether the results in Table 5 hold after controlling for earnings surprise.

Surprise is computed as the actual EPS less the forecasted EPS, scaled by the stock price at the beginning of the fourth quarter, where the forecasted EPS is the last consensus analyst forecast reported on I/B/E/S before the GCAR disclosure. As Table 2 shows, the majority of firms in the partial sample have negative earnings surprises. Untabulated comparisons of descriptive statistics show that firms in the partial sample are larger than firms in the full sample (mean total assets = $267.5 million; mean market value of equity = $74.5 million), but are similarly distressed. Consistent with the larger size, firms in the partial sample are more likely to engage Big 4 auditors (87.1 percent). They have a higher frequency of covenants restricting the firm from presenting financial statements with a GCAR (23.0 percent). The highest frequency of GCARs is observed in the year 2000, as for firms in the full sample.

The cumulative size-adjusted excess return in the three-day window (0, +2) commencing with the event date for firms in the partial sample is −8.37 percent (compared with −6.28 percent for the firms in the full sample). 72 firms with forecast data issue a press release disclosing the GCAR at least three days prior to the 10-K filing date. The cumulative three-day return for these firms is −14.34 percent (as compared to −14.64 percent for 105 early disclosers in the full sample).

We re-estimate the regressions in Table 5, adding *Surprise* to Models (1) and (2) (labeled Models 1a and 2a, respectively). Table 6 presents the regression results using the partial sample. The coefficient for *Surprise* is significant in every estimation in the expected positive direction and the R^2 increases materially. The results are otherwise very similar to those reported in Table 5, the main difference being that the control variables *LMValue* and *δIncome* are not significant for the Early Disclosers sample. The intercept is significant in Model (1)a in each case, showing a negative reaction to the GCAR announcement even after controlling for information on financial performance that may become available at the time. However, consistent with Table 5, the intercept is not significant in either estimation of Model (2)a. *CitedDebt*, *FSCovenant*, *GCCovenant*, and *InstOwn* have significant negative coefficients in Model (2)a, both for the 10-K Disclosers sample and the Early Disclosers sample.

Effect of Institutional Holdings

The results in Tables 5 and 6 suggest that the level of institutional ownership is important in explaining investors' reaction to the GCAR disclosure. We explore the issue further in this section by partitioning the sample into two groups based on the level of institutional holdings. We do a median split and form two subsamples: "low institutional ownership" firms and "high institutional ownership" firms. Although we use the label "high institutional ownership" to describe the group with larger institutional holdings, the median for the sample as a whole is only 7.0 percent (as reported in Table 2), so the threshold to enter the high institutional ownership group is relatively low by contemporary levels of institutional investment. Firms in the low institutional ownership subsample are generally smaller than firms in the large institutional ownership subsample. Low institutional holdings firms are more likely to be late in their 10-K filing and to have non-Big 4 auditors (untabulated).

When we study the daily size-adjusted excess returns around the event date for the two subsamples, it is clear that firms with high institutional holdings drive the results in Table 3. Table 7 shows the daily excess returns for the two subsamples. A significant mean excess return is evident only on Day +1 for firms with low institutional holdings, and a disproportionate number of firms in this group report a significant negative excess return only on that day. The three-day cumulative excess return, beginning with the event day, is −3.05 percent. In contrast, significant negative returns are observed on Day 0, Day +1, and Day +2 for firms with high institutional holdings. Significantly more firms report a negative excess return on each of these days, compared to the 21-day average. The three-day cumulative excess return is −9.53 percent.

We estimate Models (1) and (2) separately for the two subsamples, including both 10-K Disclosers and Early Disclosers, with the event date being the date of the first disclosure. The results are reported in Table 8, showing that the intercept is significantly negative in both the low and high institutional ownership subsamples. That is, after controlling for other financial information, there is a negative return attributable to the GCAR.

When we estimate Model (2) on the low institutional holdings subsample, none of the independent variables are significant. However, when we estimate Model (2) on the high institutional ownership subsample, the coefficients for the variables of interest (*CitedDebt*, *GCCovenant*, and *InstOwn*) are all statistically significant, as in Table 5, consistent with H2, H3, and H4, respectively. *FSCovenant* has a significantly negative coefficient, as well. The evidence suggests that

TABLE 6

OLS Regressions on the Market Reaction to the GCAR Disclosure (Controlling for Earnings Surprise)

Model 1a: $CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \beta_8 Surprise + \varepsilon$

Model 2a: $CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \beta_8 Big4 + \beta_9 CitedDebt + \beta_{10} FSCovenant + \beta_{11} GCCovenant + \beta_{12} InstOwn + \beta_{13} Surprise + \varepsilon$

		10-K Disclosers Sample (n = 425)				Early Disclosers Sample (n = 72)			
		Model (1)a		Model (2)a		Model (1)a		Model (2)a	
Variable	Exp. Sign	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept		−0.0845	−2.58*	0.0256	0.64	−0.1271	−2.35*	0.0593	0.72
LMValue	+	−0.0022	−0.28	0.0124	1.54	−0.0041	−0.23	0.0225	1.24
EBIT	+	−0.0399	−1.44	−0.0232	−0.88	−0.0458	−0.97	0.0431	0.98
δIncome	+	0.0020	0.16	0.0093	0.76	0.0594	1.81	0.0520	1.63
CFOPS	+	−0.0112	−0.35	0.0040	0.13	−0.0660	−0.97	−0.0281	−0.42
Zscore	−	−0.0056	−1.20	−0.0012	−0.27	0.0022	0.25	0.0043	0.50
ExitValue	+	−0.0079	−0.32	−0.0019	−0.08	0.0630	0.84	0.0362	0.52
LateFiler	+	0.0275	1.41	0.0229	1.23	—	—	—	—
Big4	−	—	—	−0.0486	−1.87	—	—	−0.0127	−0.23
CitedDebt	−	—	—	−0.0450	−2.38*	—	—	−0.1006	−2.19*
FSCovenant	−	—	—	−0.0542	−2.58*	—	—	−0.1181	−2.34*
GCCovenant	−	—	—	−0.0521	−2.06*	—	—	−0.1241	−2.47*
InstOwn	−	—	—	−0.2324	−4.09**	—	—	−0.2976	−2.38*
Surprise	+	0.0060	2.53*	0.0059	2.62**	0.1238	3.86**	0.0831	2.52*
R^2 (adj)		0.033		0.131		0.103		0.168	

(continued on next page)

TABLE 6 (continued)

*, ** $p < 0.05$, and $p < 0.01$, respectively. Significance levels are for two-tailed t-tests.

The dependent variable is the size-adjusted three-day (0, +2) cumulative excess return related to the GCAR disclosure date (for 10-K Disclosers, the 10-K filing date; for Early Disclosers, the GCAR-disclosing press release date).

Variable Definitions:

LMValue = natural log of market value of equity in millions;
EBIT = industry-adjusted earnings before income taxes and interest, scaled by assets;
δIncome = change in net income, scaled by assets;
CFOPS = cash flow from operations/total assets;
Zscore = a composite measure of financial distress based upon Zmijewski (1984);
ExitValue = the firm's exit value, computed based upon Berger et al. (1996);
LateFiler = 1 if the firm's 10-K was filed five days or more after the SEC statutory filing date, and 0 otherwise;
Big4 = 1 if audited by a Big 4 audit firm or one of its predecessors, and 0 otherwise;
CitedDebt = 1 if the GCAR cited a financing issue, and 0 otherwise for the 10-K sample, and 1 if the press release attributed the GCAR to a financing issue, and 0 otherwise for the Early Disclosers sample;
FSCovenant = 1 if the firm has a debt covenant that requires the client to present audited financial statements, and 0 otherwise;
GCCovenant = 1 if the firm has a debt covenant that requires the client to present an auditor's report without an explanatory paragraph noting a going concern limitation, and 0 otherwise;
InstOwn = the fraction of outstanding common stock owned by institutional investors; and
Surprise = (actual EPS − mean forecasted EPS)/stock price at beginning of the fourth quarter.

TABLE 7

Daily Excess Returns for Low and High Institutional Ownership Subsamples

	Low Institutional Ownership (n = 597)				High Institutional Ownership (n = 597)			
Day	Mean	Std. Dev.	Median	% Negative	Mean	Std. Dev.	Median	% Negative
−5	0.0043	0.1064	−0.0013	51.4%	0.0075	0.1100	−0.0025	52.8%
−4	0.0010	0.1087	−0.0003	50.4%	0.0036	0.0966	−0.0007	52.6%
−3	−0.0045	0.1132	−0.0016	51.9%	0.0011	0.0991	−0.0009	51.9%
−2	−0.0030	0.1098	−0.0009	50.8%	0.0038	0.1521	−0.0025	53.8%
−1	0.0013	0.1257	−0.0021	52.9%	−0.0006	0.1061	−0.0020	52.6%
0	−0.0086	0.1317	−0.0055	56.6%	−0.0215**	0.1281	−0.0103	60.7%**
1	−0.0158**	0.1333	−0.0080	59.0%*	−0.0540**	0.1346	−0.0296	66.3%**
2	−0.0061	0.1234	−0.0049	57.1%	−0.0198**	0.1196	−0.0082	59.1%**
3	0.0024	0.1142	−0.0021	51.8%	0.0060	0.1247	−0.0012	54.9%
4	−0.0022	0.1331	−0.0022	54.6%	0.0120	0.1255	0.0001	50.2%
5	0.0020	0.1151	−0.0028	51.4%	−0.0012	0.1117	−0.0043	54.2%

*, ** $p < 0.05$, and $p < 0.01$, respectively, two-tailed.

The event date is the GCAR disclosure date (the earlier of a GCAR-disclosing press release or the 10-K filing).

The table presents statistics for the daily size-adjusted excess return. Significance levels are for t-tests for differences in the mean daily excess return from zero or for Chi-square tests for differences in the percentage of negative excess returns in a given day from the average for the 21 days around the event date for sample firms.

TABLE 8

OLS Regressions on the Market Reaction to the GCAR Disclosure for Low and High Institutional Ownership Subsamples

Model 1: $CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \varepsilon$

Model 2: $CAR = \alpha + \beta_1 LMValue + \beta_2 EBIT + \beta_3 \delta Income + \beta_4 CFOPS + \beta_5 Zscore + \beta_6 ExitValue + \beta_7 LateFiler + \beta_8 Big4 + \beta_9 CitedDebt + \beta_{10} FSCovenant + \beta_{11} GCCovenant + \beta_{12} InstOwn + \varepsilon$

		Low Institutional Ownership (n = 597)				High Institutional Ownership (n = 597)			
		Model (1)		Model (2)		Model (1)		Model (2)	
Variable	Exp. Sign	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
Intercept		−0.0578	−2.79**	−0.0317	−1.38	−0.0836	−3.37**	−0.0304	−0.97
LMValue	+	0.0060	0.97	0.0080	1.29	0.0005	0.01	0.0100	1.57
EBIT	+	−0.0142	−1.14	−0.0107	−0.85	0.0040	0.15	0.0175	0.69
δIncome	+	0.0089	1.02	0.0089	1.02	0.0227	1.97*	0.0182	1.63
CFOPS	+	−0.0033	−0.21	−0.0008	−0.05	−0.0272	−0.91	−0.0104	−0.36
Zscore	−	−0.0021	−0.71	−0.0011	−0.38	−0.0028	−0.71	0.0017	0.45
ExitValue	+	0.0293	0.32	0.0161	0.18	0.0195	0.96	0.0213	1.09
LateFiler	+	0.0113	0.83	0.0071	0.51	0.0152	0.93	0.0253	1.60
Big4	−	—	—	−0.0039	−0.28	—	—	−0.0510	−2.41*
CitedDebt	−	—	—	−0.0173	−1.18	—	—	−0.0464	−2.84**
FSCovenant	−	—	—	−0.0193	−1.38	—	—	−0.0458	−2.58*
GCCovenant	−	—	—	−0.0408	−1.70	—	—	−0.0498	−2.36*
InstOwn	−	—	—	−0.2753	−0.90	—	—	−0.1857	−3.62**
R^2 (adj)		0.007		0.025		0.012		0.076	

(continued on next page)

TABLE 8 (continued)

*, ** $p < 0.05$, and $p < 0.01$, respectively. Significance levels are for two-tailed t-tests.

The dependent variable is the size-adjusted three-day (0, +2) cumulative excess return related to the GCAR disclosure date (the earlier of a GCAR-disclosing press release or the 10-K filing).

Variable Definitions:

LMValue = natural log of market value of equity in millions;
EBIT = industry-adjusted earnings before income taxes and interest, scaled by assets;
δIncome = change in net income, scaled by assets;
CFOPS = cash flow from operations/total assets;
Zscore = a composite measure of financial distress based upon Zmijewski (1984);
ExitValue = the firm's exit value, computed based upon Berger et al. (1996);
LateFiler = 1 if the firm's 10-K was filed five days or more after the SEC statutory filing date, and 0 otherwise;
Big4 = 1 if audited by a Big 4 audit firm or one of its predecessors, and 0 otherwise;
CitedDebt = 1 if the GCAR cited a financing issue, and 0 otherwise;
FSCovenant = 1 if the firm has a debt covenant that requires the client to present audited financial statements, and 0 otherwise;
GCCovenant = 1 if the firm has a debt covenant that requires the client to present an auditor's report without an explanatory paragraph noting a going concern limitation, and 0 otherwise; and
InstOwn = the fraction of outstanding common stock owned by institutional investors.

sophisticated institutional investors may be reacting to the auditor's assessment of the firms' financing problems, as well as to additional costs imposed because the GCAR violates a debt covenant. Interestingly, the coefficient for *Big4* is negative and significant. Recall that this coefficient is not significant when the regression is estimated on the full sample in Table 5. GCARs issued by Big 4 auditors are likely viewed as more adverse than GCARs issued by non-Big 4 auditors in the high institutional ownership subsample.

The results reported in Table 8 for the low institutional holdings sample indicate that there is some negative reaction to the GCAR announcement at low levels of institutional ownership. These results are sensitive to the cutoff employed to determine low holdings. When low institutional holdings are defined as the bottom tertile of the sample (<1.5 percent of the firm's outstanding shares) or the bottom quartile (<1.2 percent), the intercept is not significant ($p = 0.26$ and $p = 0.20$, respectively), indicating that there is no reaction to the GCAR in firms with such low levels of institutional ownership. Overall, these results suggest that institutional investors may be largely responsible for the negative reaction to GCARs that we observe.

Institutional Holdings before and after GCAR

The analysis in the preceding subsection shows that the market reaction to the GCAR is associated with the level of institutional holdings in the firm. The next analysis provides additional evidence that the GCAR affects institutional holders' investment decisions. This analysis follows the work of Kausar et al. (2006), who find a decline in net institutional holdings in firms that receive a GCAR. We extend Kausar et al.'s (2006) study to determine whether the decline in institutional holdings depends upon the characteristics of the GCAR. Recall that institutional holdings data are obtained originally from quarterly 13F filings. We compare the institutional holdings for the firm at two points in time: the quarter-end that precedes the calendar quarter in which the GCAR is disclosed, and the end of the first full quarter after the GCAR is disclosed.[30] Table 9 shows that the mean institutional ownership drops from 13.3 percent in the quarter-end preceding the GCAR to 8.6 percent at the end of the quarter following the GCAR. The decline is statistically significant, and confirms the result reported by Kausar et al. (2006).

We partition the sample by whether the GCAR cites a financing problem. The 424 firms that are issued GCARs citing a financing problem have a higher level of institutional ownership before the GCAR, on average (14.7 percent, versus 12.5 percent for firms where the GCAR does not cite a financing problem), but they experience a significantly greater decline in institutional ownership. We partition the sample also according to whether the firm has a covenant restricting the GCAR. The level of institutional ownership for the firms in this group (mean = 21.2 percent) is initially much higher than for firms without such a covenant (mean = 11.8 percent). There is no statistical difference in the level of institutional ownership following the GCAR because of a steep drop in the institutional ownership in firms where the GCAR violates a covenant.

In untabulated results, we also estimate a regression using the change in institutional ownership as the dependent variable and *GCCovenant* (the presence of a covenant restricting the GCAR) and *CitedDebt* (the citing of financing problems in the GCAR) as the main independent variables. We control for size (*LMValue*), financial performance (*EBIT*, *δIncome*, *CFOPS*, *Zscore*, and *ExitValue*), auditor type (*Big4*), and whether there is a covenant requiring financial statements (*FSCovenant*). Of the variables of interest, only the coefficient for *GCCovenant* is statistically significant. It has a negative sign, consistent with the univariate analysis in Table 9.

[30] For example, for a firm disclosing a GCAR on March 15, 2005, we would study the change in institutional ownership from December 31, 2004 to June 30, 2005.

TABLE 9

Changes in Institutional Ownership Following GCAR

	n	Mean Institutional Ownership before GCAR	Mean Institutional Ownership after GCAR	Difference
All Firms	1,194	0.133	0.086	−0.046
t-statistic				5.68 **
Reasons Cited for GCAR				
GCARs citing financing problems	424	0.147	0.084	−0.063
GCARs not citing financing problems	770	0.125	0.088	−0.037
Test of differences	1,194	2.33 *	0.50	3.13 **
Covenant Restricting the GCAR				
Firms with covenant restricting GCAR	193	0.212	0.095	−0.117
Firms without covenant restricting GCAR	1,001	0.118	0.085	−0.033
Test of differences	1,194	5.70 **	0.79	5.66 **
Institutional Ownership				
Firms with low institutional ownership	597	0.021	0.032	0.011
Firms with high institutional ownership	597	0.244	0.141	−0.103
Test of differences	1,194	34.50 **	14.83 **	15.37 **

*, ** $p < 0.05$, and $p < 0.01$, respectively, two-tailed.

The table shows the mean level of institutional ownership as a proportion of total common shares outstanding in firms before and after they are given a GCAR. t-statistics are for tests for differences in means before and after the GCAR in the case of the full sample, and for differences between subgroup means in other cases.

These results are driven by firms with higher levels of institutional ownership. Table 9 shows that when we do a median split of the sample along institutional ownership, as in Tables 7 and 8, there is little change in the mean institutional ownership following the GCAR in the low institutional ownership subsample, but a substantial drop (from 24.4 percent to 14.1 percent) in the high ownership subsample. We estimate a regression with the change in institutional holdings as the dependent variable in each subsample. Results for the high institutional ownership subsample are similar to the results reported for the sample as a whole, but in the low institutional subsample neither *GCCovenant* nor *CitedDebt* have significant coefficients ($p = 0.48$ and 0.60, respectively).

Table 9 does not directly test institutional selling in response to the GCAR, because the change in institutional ownership levels around the time the GCAR is issued does not necessarily mean that institutional owners reacted to the GCAR. Still, the evidence is consistent with institutions being net sellers of stock in GCAR-receiving firms, and in particular, in firms receiving GCARs that cite financing problems and in firms that have a debt covenant that restricts them from receiving a GCAR.

IV. ADDITIONAL ANALYSES AND ROBUSTNESS TESTS

GCAR Type I Errors

It has long been speculated that a GCAR may increase the likelihood of bankruptcy. Prior

studies have referred to this as the "self-fulfilling prophecy" effect of the GCAR (e.g., Mutchler 1984; Louwers et al. 1999). Several studies have investigated the GCAR's accuracy in predicting the likelihood that a firm will go bankrupt (e.g., Geiger and Rama 2006). We extend this literature by considering whether auditors are less likely to make Type I errors (issuing GCARs to firms that survive for more than one year) for GCARs with certain characteristics than for others. We conduct this analysis to rule out the possibility that the more adverse market reaction observed for GCARs citing financing problems and GCARs violating covenants is due to these firms having a greater likelihood of going bankrupt than other firms receiving GCARs.

We begin by determining the post-GCAR status of the companies in the sample. We find that 174 of the 1,194 firms file for bankruptcy (Chapter 7 or 11) at the federal or state level in the 12-month period following the date of the financial statements for which the auditor issued a GCAR, a Type I error rate of 85.4 percent. To determine whether Type I errors are lower for certain types of GCARs, we estimate a logistic model where the dependent variable, *Bankrupt* (1 if bankrupt, and 0 otherwise) indicates whether the firm goes bankrupt in the 12 months following the GCAR. *CitedDebt* and *GCCovenant* are the variables of interest in this analysis. We use *FSCovenant* to ensure that *GCCovenant* is not simply picking up firms with covenants requiring audited financial statements. As control variables we use *LMValue*, since larger firms may have greater ability to survive, *EBIT*, *δIncome*, *CFOPS*, *Zscore*, *ExitValue*, and *LateFiler* to represent financial condition, and *Big4*, since Big 4 auditors may have superior accuracy in issuing GCARs. The variables of interest are not significant (lowest $p = 0.16$).

Another potential adverse economic consequence a firm faces is delisting. As many as 467 of the 1,194 firms in our sample (39.1 percent) are delisted during the 12 months following the date of the financial statements carrying the GCAR. We test whether the likelihood of delisting depends upon the characteristics of the GCAR. To do this, we replace the dependent variable in the bankruptcy model described above with *Delist*, where *Delist* equals 1 if the firm is delisted from the stock exchange in the 12 months following the GCAR. Again, the GCAR-related variables do not have significant coefficients (lowest $p = 0.24$). The likelihood that a firm will go bankrupt or delist is not greater for firms with GCARs citing financing problems or firms with GCARs that violate debt covenants.

Potentially Confounding Disclosures

Given the nature of the firms in the sample and their financial condition at the time of the GCAR disclosure, it is plausible that the negative reaction we detect is related to other significant firm-specific events that are concurrently disclosed to investors. To rule out this potential explanation, we review all 8-K filings made by each firm in the sample during the period starting five days before the event date and ending five days after. We retain in the sample those firms that file 8-Ks, reproducing previously disseminated earnings or GCAR announcements, but drop 147 firms reporting other events. The principal events reported in the filings by omitted firms are financing issues, purchase or sale of assets, and manager or director changes. We re-estimate Models (1) and (2) on a sample that excludes these firms. The intercept is statistically significant in Model (1) for the 10-K Disclosers sample. The coefficients for the variables of interest in Model (2) are significant at previous levels, as shown in Table 5.

We also review the Early Disclosers sample to identify significant disclosures of other events. We identify 10 firms with significant events that were not previously disclosed (3 executive turnovers, 2 litigation events, 3 loan defaults, and 2 mergers). We delete these firms from the Early Disclosers sample and re-estimate regressions in Table 5 for the remaining 95 firms. The intercept is statistically significant in Model (1) and the coefficients for the variables of interest in Model (2) are significant at previous levels.

GCAR-Related Variables and Returns at Earnings Announcement

A potential concern is that the GCAR-related variables (*GCCovenant* and *CitedDebt*) may proxy for factors that are associated with the firm's financial condition and that we have not adequately captured with our control variables. We address this concern by estimating Models (2) and 2a at another significant event date when the GCAR should not be relevant. The sample for this analysis consists of 349 firms in the full sample of firms (191 firms in the partial sample) that first announce the GCAR at the 10-K date, but that previously issue earnings announcements. The event date is the earnings announcement date. Since the GCAR is not known at this time, the GCAR-related variables should be associated with excess returns in the event window (0, +2) only if they are surrogates for financial distress. The only statistically significant coefficients when we estimate Models (2) and 2a are for the control variables representing size and earnings surprise (Model 2: *LMValue*; Model (2)a: *LMValue* and *Surprise*). It seems unlikely that *GCCovenant* and *CitedDebt* represent other factors associated with financial distress.

Additional Variables

Recall that we expect investors to react adversely to GCARs that cite financing problems. We do not expect the other commonly cited reasons (see Table 4, Panel A) to be incrementally informative because the cited circumstances are likely to be known to investors. Still, we individually test to see whether each reason is informative, using a regression in the form of Model (2). We do not find an incremental negative effect on returns for any reason except financing problems (lowest $p = 0.13$).

Table 4, Panel B indicates that there are 92 firms with covenants that require that the audit report be "unqualified." We do not include these with the 195 firms carrying covenants that prohibit a GCAR in our main tests because auditing standards treat the GCAR as an unqualified audit report. However, it is possible that a lender that prohibits a qualified opinion intends to prohibit a GCAR, since the GCAR is sometimes referred to as a "going concern qualification." When we add a variable to Model (2) to indicate whether the firm has a covenant requiring an unqualified report, the coefficient is not significant ($p = 0.35$), suggesting that investors do not react more adversely to the GCAR announcement if such a covenant is present.

We include several additional variables in Models (1) and (2) as potential explanatory variables for the event period returns. *Restate* incorporates adverse expectations that investors may have for restating firms. *PreReturn* captures pre-event returns and is used as a proxy for investors' anticipation of the GCAR. We use *Bankrupt* since investors may make a distinction at the time of the GCAR announcement between firms that will imminently go bankrupt and those that will not (Blay and Geiger 2001). Similarly, *Delist* is included because investors may anticipate that additional costs will be borne by firms that are likely to be delisted.[31] The coefficients for *Restate*, *PreReturn*, *Bankrupt*, and *Delist* are not significant at the thresholds we use in this paper in any of the model-sample combinations used in Table 5 (lowest $p = 0.10$ for *Restate*, in Model (2) for both samples). The variables of interest continue to be significant at the same levels as in Table 5.

In order to see whether there is a difference in the market reaction to the GCAR in the pre- and post-Regulation FD periods, we use a variable *PostFD*, coded as 1 for all GCARs with event dates following the effective date of Regulation FD. We also use interaction terms *PostFD* * *CitedDebt*, *PostFD* * *GCCovenant*, and *PostFD* * *InstOwn*. This analysis is motivated by the information advantage argument that suggests that in the period before Regulation FD, institu-

[31] *Restate* equals 1 if the firm reports a restatement in the 12 months preceding the event date, and 0 otherwise; *PreReturn* is measured as the stock return over the period starting from the end of the fiscal year for which the GCAR is obtained and culminating two days prior to the event date; *Bankrupt* is as defined earlier; and *Delist* is as defined earlier.

tional owners may react to the information in the GCAR before, rather than during, the event date because of access to predisclosure information. We re-estimate Models (1) and (2). The intercept in Model (1) and the independent variables of interest in Model (2) are significant at previous levels. The additional variable and interaction terms used are not significant (lowest $p = 0.26$), indicating there is no difference in the reaction of individual investors at the announcement date before and after Regulation FD.

Alternative Measurements of Variables

For 235 firms in the partial sample, there is no prior press release announcing earnings, and the 10-K provides the first news on earnings. For the remaining firms, there is a prior earnings release, and earnings are no longer a surprise. For these firms, *Surprise* arguably should be coded as 0. Treating *Surprise* as 0 for these firms does not change the statistical significance of *Surprise* or alter other results in Table 6.

Our tests use a five-day grace period for *LateFiler*. In robustness tests, we classify a firm as a late filer if it is one, two, three, or four days late in filing. We also use two alternative measures of financial distress to replace the Zmijewski (1984) distress metric: Altman's Z-score (Altman 1968) and a metric based on the Hillegeist et al. (2004) financial distress model. Our results are robust to the use of these alternative measures (i.e., coefficients for the variables of interest continue to be significant at reported levels).

We repeat all our tests using buy-and-hold abnormal returns (BHAR) rather than cumulative abnormal returns (CAR) and obtain essentially similar results. The coefficients for the variables of interest are significant at reported levels with one exception: in Table 5, in the 10-K Disclosers sample, the coefficient for *GCCovenant* is significant at a lower level ($p = 0.06$) than reported previously. Since this is a short-window study, we choose to report CAR in the tables to be consistent with other short-window studies, particularly prior GCAR event studies.

V. CONCLUSIONS

There have been questions raised in the past about whether auditors should be required to make going concern judgments. One view has been that GCARs do not provide investors with new information, since investors can make going concern assessments adequately using other sources of information. Others have argued that the auditor's access to private information regarding the client makes the GCAR informative. Prior research that attempts to ascertain the usefulness of the GCAR by studying the market reaction to the GCAR announcement provides mixed results. The sample in this study is substantially larger than samples used previously and covers many years. We find that investors react to the auditor's assessment and adjust their valuations accordingly. In addition, we show that the market reaction to the GCAR depends on certain GCAR attributes. Mixed results in prior studies could be due to differences in sample composition with respect to the reasons cited by the auditor for the GCAR, the presence of covenants prohibiting the GCAR, and the level of institutional ownership.

Our results suggest that the informativeness of the GCAR depends upon its content. Auditors generally cite a reason for issuing the GCAR. Very often, the cited reason is the firm's distressed financial condition, which may not add to what investors can already glean from the financial statements. We find that the reaction is more adverse for firms when the auditor attributes the GCAR to problems related to obtaining financing.

It has been widely speculated that GCARs may have a self-fulfilling prophecy effect; that is, the auditor's decision to issue a GCAR results in firms finding it difficult to obtain capital and to carry on business as usual and, that as a consequence, the GCAR leads to failure. Our study shows one specific way in which the GCAR affects contracting costs. A significant number of firms have lending agreements that include covenants that specifically require that the firm not obtain a

GCAR. Consistent with other studies that find that investors are sensitive to the violation of debt covenants, investors react more negatively when the firm violates a covenant requiring GCAR-free financial statements. However, neither the auditor's attribution of the GCAR to financing problems nor the violation of a covenant requiring GCAR-free financial statements increases the likelihood that the firm will go bankrupt or delist from the stock exchange, relative to other firms receiving GCARs. In reacting negatively to these two types of GCARs, investors are recognizing potential increases in the cost of obtaining capital rather than to imminent bankruptcy or delisting.

Finally, we find that the reaction to the GCAR is more negative when the level of institutional holdings in the firm is high. Interestingly, there is no reaction at all to the GCAR in firms with very low levels of institutional ownership. At higher levels of institutional ownership, the magnitude of negative reaction increases for firms where the GCAR cites financing problems and firms for which the GCAR results in a debt covenant violation. Also, when we study the level of institutional ownership before and after the GCAR, we find significant net institutional selling in these types of firms. We attribute these results to sophisticated investors' awareness of the firm's financing situation and the covenants carried by the firm's debt.

REFERENCES

Alford, A., J. Jones, and M. Zmijewski. 1994. Extension and violation of the statutory SEC form 10-K filing requirements. *Journal of Accounting and Economics* 17 (January): 229–254.

Altman, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23 (September): 589–609.

American Institute of Certified Public Accountants (AICPA). 1988. *The Auditor's Consideration of an Entity's Ability to Continue as a Going Concern.* Statement on Auditing Standards No. 59. New York, NY: AICPA.

Banz, R. W. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9 (March): 3–18.

Barth, M. E., W. H. Beaver, and W. R. Landsman. 1998. Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics* 25 (February): 1–34.

Bellovary, J. L., D. E. Giacomino, and M. D. Akers. 2006. Weighing the public interest: Is the going concern opinion still relevant? *The CPA Journal* 76 (January): 16–21.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics.* New York, NY: John Wiley.

Beneish, M. D., and E. Press. 1993. Costs of technical violation of accounting-based debt covenants. *The Accounting Review* 68 (April): 233–257.

———, and ———. 1995. The resolution of technical default. *The Accounting Review* 70 (April): 337–353.

Berger, P. G., E. Ofek, and I. Swary. 1996. Investor evaluation of the abandonment option. *Journal of Financial Economics* 42 (October): 257–287.

Blay, A. D., and M. A. Geiger. 2001. Market expectations for first-time going concern recipients. *Journal of Accounting, Auditing & Finance* 16 (Summer): 209–226.

Business Recorder. 2002. Standard & Poor's Cuts Gentek's Corporate Credit Rating. (March 30).

Carrns, A. 2000. Charting one internet patient's rapid decline. *Wall Street Journal* (April 6): B1.

Chen, K. C. W., and B. K. Church. 1996. Going concern opinions and the market's reaction to bankruptcy filings. *The Accounting Review* 71 (January): 117–128.

Collins, D. W., M. Pincus, and H. Xie. 1999. Equity valuation and negative earnings: The role of book value of equity. *The Accounting Review* 74 (January): 29–61.

Dodd, P., N. Dopuch, R. Holthausen, and R. Leftwich. 1984. Qualified audit opinions and stock prices. *Journal of Accounting and Economics* 6 (April): 3–38.

Dopuch, N., R. Holthausen, and R. Leftwich. 1986. Abnormal stock returns associated with media disclosures of "subject to" qualified audit opinions. *Journal of Accounting and Economics* 8 (June): 93–117.

Ellingsen, J. E., K. Pany, and P. Fagan. 1989. SAS No. 59: How to evaluate going concern. *Journal of Accountancy* 165 (January): 24–31.

Elliott, J. A. 1982. "Subject to" audit opinions and abnormal security returns—Outcomes and ambiguities. *Journal of Accounting Research* 20 (Autumn): 617–638.

Fleak, S. K., and E. R. Wilson. 1994. The incremental information content of the going-concern audit opinion. *Journal of Accounting, Auditing & Finance* 9 (Winter): 149–166.

Geiger, M. A., and D. Rama. 2006. Audit firm size and going-concern reporting accuracy. *Accounting Horizons* 20 (March): 1–17.

Herbohn, K., V. Ragunathan, and R. Garsden. 2007. The horse has bolted: Revisiting the market reaction to going concern modifications of audit reports. *Accounting and Finance* 47 (September): 473–493.

Hillegeist, S. A., E. K. Keating, D. P. Cram, and K. G. Lundstedt. 2004. Assessing the probability of bankruptcy. *Review of Accounting Studies* 9 (March): 5–34.

Holder-Webb, L. M., and M. S. Wilkins. 2000. The incremental information content of SAS No. 59 going-concern opinions. *Journal of Accounting Research* 38 (Spring): 209–219.

Hopwood, W., J. McKeown, and J. Mutchler. 1989. A test of the incremental explanatory power of opinions qualified for consistency and uncertainty. *The Accounting Review* 64 (January): 28–48.

Jones, F. L. 1996. The information content of the auditor's going concern evaluation. *Journal of Accounting and Public Policy* 15 (Spring): 1–27.

Kausar, A., R. J. Taffler, and C. Tan. 2006. Anomalous price reaction and differential stockholder response to going-concern audit opinions and withdrawals. Available at SSRN: http://ssrn.com/abstract=889282.

———, ———, and ———. 2009. The going-concern market anomaly. *Journal of Accounting Research* 47 (March): 212–239.

Landler, M. 2004. For Spiegel investor, a hard lesson in securities law. *The New York Times* (September 11): C1.

Louwers, T. J., F. M. Messina, and M. D. Richard. 1999. The auditor's going-concern disclosure as a self-fulfilling prophecy: A discrete-time survival analysis. *Decision Sciences* 30 (Summer): 805–824.

Mutchler, J. F. 1984. Auditor's perceptions of the going-concern opinion decision. *Auditing: A Journal of Practice & Theory* 3 (Spring): 17–30.

Norris, F. 2003. KPMG faulted on handling of Polaroid financial reports. *The New York Times* (August 23): C1, C3.

Ogneva, M., and K. R. Subramanyam. 2007. Does the stock market under-react to going concern opinions? The evidence from the U. S. and Australia. *Journal of Accounting and Economics* 43 (July): 439–452.

Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32 (Spring): 38–60.

Taffler, R. J., J. Lu, and A. Kausar. 2004. In denial? Stock market under-reaction to going-concern audit report disclosures. *Journal of Accounting and Economics* 38 (December): 263–296.

Utama, S., and W. M. Cready. 1997. Institutional ownership, differential predisclosure precision and trading volume at announcement dates. *Journal of Accounting and Economics* 24 (December): 129–150.

Walther, B. 1997. Investor sophistication and market earnings estimates. *Journal of Accounting Research* 35 (Autumn): 157–179.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22 (Supplement): 59–82.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 2107–2143

American Accounting Association
DOI: 10.2308/accr.00000001

The Effects of Reporting Complexity on Small and Large Investor Trading

Brian P. Miller
Indiana University

ABSTRACT: This study examines the effects of financial reporting complexity on investors' trading behavior. I find that more complex (longer and less readable) filings are associated with lower overall trading, and that this relationship appears due to a reduction in small investors' trading activity. Additional evidence suggests that the association between report complexity and lower abnormal trading is driven by both cross-sectional variation in firms' disclosure attributes and variations in disclosure complexity over time. Given regulatory concerns over plain English disclosures and the trend toward more disclosure, my investigation into the effects of reporting complexity on small and large investors should be of interest to regulators concerned with reporting clarity and leveling the playing field across classes of investors.

Keywords: *disclosure; readability; plain English; trading volume; small and large investors.*

JEL Classifications: *G30; G12; M41; M48.*

I. INTRODUCTION

The securities industry has historically been driven by the fundamental concept that "all investors, whether large institutions or private individuals, should have access to certain basic facts about an investment prior to buying it, and so long as they hold it" (SEC 2010, under "Introduction"). Consistent with this philosophy, investors are increasingly able to access

This study is based on my dissertation, completed at The Pennsylvania State University. I am grateful for the guidance and assistance of my dissertation committee: Paul Fischer (co-chair), Andy Leone, Karl Muller (co-chair), and Arun Upneja. Useful comments and discussions have also been provided by Walid Al-Issa, Orie Barron, Daniel Beneish, Neil Bhattacharya, Christine Cheng, Mark Evans, Bob Jennings, Dan Givoly, Karen Hennes, Max Hewitt, Steven Huddart, Steven Kachelmeier (senior editor), Ben Lansford, Feng Li, Laureen Maines, Jim Vincent, Teri Lombardi Yohn, Hal White, two anonymous referees and seminar participants at Carnegie Mellon University, Duke University, Indiana University, Michigan State University, The Ohio State University, The Pennsylvania State University, Purdue University, Southern Methodist University, University of Michigan, The University of North Carolina at Chapel Hill, University of Notre Dame, University of Rochester, University of Toronto, and Washington University in St. Louis. I also express gratitude to Justin Patterson and StyleWriter software engineers for programming assistance. Finally, I acknowledge financial support from the Smeal Dissertation Research Award.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: December 2008
Accepted: May 2010
Published Online: November 2010

more data from longer public filings.[1] However, this trend may not necessarily aid investors if it is more costly to extract useful information from larger and more complex disclosures (Bloomfield 2002).

Although there is evidence regarding the effects of reporting length on long-term market prices (You and Zhang 2009), the effects on individual market participants and aggregate trading volume remain unexamined. As such, I investigate the effects of reporting complexity (length and readability) on small and large investors' trading behavior (volume and consensus) around 10-K filings. This investigation should interest regulators concerned with both reporting clarity and leveling the playing field across classes of investors.[2]

I first consider the effect of more complex reports on abnormal trading around the 10-K filing. I hypothesize that, faced with more complex reports, investors may elect not to process the report because doing so is too costly. Consequently, these investors will not initiate trades in response to the report (Grossman and Stiglitz 1980; Bloomfield 2002). After controlling for information content, I predict that more complex filings will lead to lower trading volume, and these effects will be most pronounced among smaller investors (due to their limited processing capabilities and expertise). Although theory predicts that higher processing costs affect investor trading behavior, *a priori* it is unclear whether the trend toward more complex filings will affect overall investor trading behavior. First, contemporaneous advances in technology are likely to dampen overall processing costs, making it easier for all investors to process more complex reports. Second, the increased availability of other information sources (e.g., media and analyst reports) could mitigate the effects of these attributes on investor trading behavior.[3]

I examine approximately 13,000 10-K reports filed between 1995 and 2006 and find evidence consistent with more complex reports affecting trade behavior around the filing. Specifically, I find that (1) more complex reports are associated with lower levels of aggregate trading volume; and (2) the relationship between report complexity and trade activity appears to be driven by a reduction in small investor trade volume. These results are robust to controlling for the information content of the report and factors such as profitability and persistence that have been shown to vary with reporting complexity. Overall, the evidence is consistent with more complex reports being more costly to process, with the effects being most pronounced among small traders.

In addition to trading volume, I also investigate whether more complex reports lead to more dispersion in the inferences drawn by those investors who do elect to trade. Based on prior experimental evidence (Barron et al. 2004), I predict that when reports are more complex, small investors will interpret the information in a more diverse manner. Consistent with this prediction, I find that more complex reports are associated with a decrease in consensus among small investors, but not among large investors.

Admittedly, reporting complexity is a broad notion that captures both differences in relative report complexity across firms as well as variations in disclosure attributes over time. I attempt to isolate these potentially different effects by performing both within-year and within-firm analyses.

---

[1] Radin (2007) points out that increases in report length can be traced to new disclosures devoted to risk, compensation, pension accounting, and stock options. It is also plausible that firms have voluntarily increased disclosure over time. Bloomfield (2008) provides an excellent review of the potential theories why managers may alter their reporting in certain settings. The focus of this study is how report complexity affects different investors' trade behavior and not necessarily the causes of increased disclosure over time.

[2] SEC concerns over report clarity are highlighted in the Plain English Handbook (SEC 1998), while concerns over a level playing field are the focus of recent regulatory initiatives (e.g., limit disclosure practices that discriminate across classes of investors (SEC 2000a) and regulate the fair execution of trades (SEC 2000b).

[3] Advances in technology are likely to assist all investors with accessing and processing more complex reports. However, large investors may be more likely to take advantage of new parsing tools (e.g., Perl) to analyze reports, whereas less sophisticated investors may rely more on public information from sources such as Yahoo! Finance. These differences could further exacerbate the processing differential between small and large investors.

The results show the association between more complex reports and lower abnormal trading appears to be associated with both cross-sectional variation in firms' disclosure attributes (e.g., operational complexity) in a given year as well as variations in disclosure length over time (e.g., increased disclosure due to regulation) for a given firm.

While the evidence in this study substantiates some regulatory concerns regarding the length and the readability of mandatory filings (SEC 1998), it is unclear *a priori* which of these complexity attributes have a larger effect on investors. Hence, in addition to investigating the effects of length and readability separately, I also investigate the effects of both measures when included in the same model. I find that when both measures are analyzed simultaneously, the effects of longer reports remain significant, while the readability measures become insignificant. This finding suggests the measures are substitutes, but the effects of longer reports dominate the readability of the filings.

The evidence in this study is generally consistent with the theory of more complex filings being too costly for small investors to process in the short window surrounding the filing date. This evidence should prove useful to both academics and regulators. The findings specifically challenge some long-standing regulatory assumptions that requiring more disclosure will not only aid investors in their trading decisions, but also help level the playing field between small and large investors. Thus, although the SEC has made progress in assisting the small investor in gaining electronic access to financial information (Asthana et al. 2004), simply making more data available may not benefit small investors if the reports are too long and complex to process.

Section II provides additional background and hypothesis development. Section III outlines the research design. Sample selection and descriptive data are provided in Section IV. Empirical results are summarized in Section V. Section VI provides additional analyses, and Section VII concludes.

II. BACKGROUND AND HYPOTHESIS DEVELPMENT

Efforts by the SEC to make financial data more readable and understandable date back to the Securities Act of 1933 (Firtel 1999). In 1969, the SEC released the *Wheat Report*, which indicated that the length and complexity of prospectuses prevented the average investor from readily understanding them. The report went on to recommend against unnecessarily long, complex and/or verbose writing. These concerns over lengthy reports were reiterated by Arthur Levitt in his 1997 remarks to the Securities Regulation Institute, when he asserted that "[i]n many cases, the problem is not a lack of information; quite the opposite. Too much information can be as much a problem as too little. More disclosure does not always mean better disclosure" (Levitt 1997, under "More Timely and Useful Disclosure for Investors").

Under Levitt's leadership, the SEC adopted the 1998 plain English regulation, SEC Rule 421(d), which required issuers to use plain English principles in the design of prospectuses (SEC 1998). Despite this regulation and concerns that financial reports have simply become too long for investors to process (Paredes 2003), there has been a substantial increase in report length over the past decade (Li 2008). Whether this additional data is beneficial to investors is unclear. In fact, some accounting practitioners argue that the useful information disclosed in annual filings is now hidden among a plethora of boilerplate, redundant, immaterial, or even irrelevant data, making disclosures increasingly difficult to process (Radin 2007).

Related Literature

Despite the importance practitioners and regulators have placed on disclosure clarity and length, there is little large-scale empirical evidence of effects of these attributes on financial statement accessibility. Most of the early work in this area investigates the readability of financial

statements and footnotes in small sample sizes. Jones and Shoemaker (1994) summarize this literature, concluding that financial disclosures are difficult to read and "inaccessible" to a large proportion of unsophisticated investors.

Recent readability research focuses primarily on managerial incentives to disclose more or less readable reports under different circumstances. Li (2008) documents an association between length and readability of 10-K filings and both profitability and earnings persistence, while Nelson and Pritchard (2007) find that firms subject to greater shareholder litigation use more readable language in their disclosures and avoid boilerplate warnings. Both studies focus on the managerial discretion in reporting attributes, but neither study links the effects of reporting complexity to investor behavior.

Another line of research investigates the market reaction to report complexity. You and Zhang (2009) document that the market under-reaction to 10-K filings is stronger for firms with longer 10-K filings. Other linguistics research on financial disclosures investigates the market response to optimistic and pessimistic tone in media (Tetlock 2007; Tetlock et al. 2008) and earnings press releases (Davis et al. 2007). Li (2009) finds that MD&A tone and accruals predict future earnings and stock returns. Overall, research on length, readability, and tonality focuses on market responses, but does not address whether these reporting attributes affect the specific trade behavior of large and small investors.

A concurrent working paper by Loughran and McDonald (2010) also investigates report complexity of 10-K filings. As discussed in greater detail in Section VI, there are several differences in the research design and readability proxies between this study and Loughran and McDonald (2010). Despite these differences, both studies provide evidence that improvements in report readability lead to increased trading activity among small investors. Further, both studies provide evidence that the effects are most evident when readability is measured using proxies based on the SEC's plain English guidelines. The two studies differ in that Loughran and McDonald (2010) focus primarily on report readability, while the focus of the current study encompasses the effects of both length and readability.[4] This distinction is most apparent in the evidence provided by this study that the effects of report readability are subsumed by the effects of length when both measures are simultaneously examined. Therefore, this study suggests that the effects of length on investor trading behavior are more important than the effects of readability.

Finally, in addition to the direct contributions to the emerging literature on reporting complexity, this study also adds to several large bodies of research that (1) demonstrate the benefits of more informative disclosures (Lang and Lundholm 1996; Botosan 1997; Francis et al. 2002), (2) examine price and trading volume reactions to information releases (Beaver 1968; Morse 1981; Cready and Mynatt 1991; Asthana and Balsam 2001; Griffin 2003; Li and Ramesh 2009), and (3) document the differential trading behavior of small and large investors to various information events (Bhattacharya 2001; Shanthikumar 2004; Asthana et al. 2004; Allee et al. 2007; Bhattacharya et al. 2007; Mikhail et al. 2007). Although this literature provides a necessary background for the research conducted in this study, it does not provide any direct evidence on the effects of report complexity on different classes of investors.

Hypothesis Development

Trading Volume

Consistent with evidence that trading volume is the most visible indicator of investors' response to public disclosures (Cready and Hurtt 2002), I focus on abnormal trading activity around

[4] Loughran and McDonald (2010) do include a measure of length in their regressions. However, the authors investigate the effect of a change in readability on changes in abnormal trading and include the level (not change) in length. Thus, any deviations in length over time are not directly examined in their analysis.

the 10-K filing to assess investors' response to variations in report complexity. I hypothesize that faced with more complex filings, some investors will elect not to process the report because doing so is too costly. Consequently, these investors will not initiate trades in response to the report (Grossman and Stiglitz 1980; Bloomfield 2002). After controlling for the information content in the report, I expect that more complex filings will be associated with lower total trading volume. Formally, the first hypothesis (in alternative form) is as follows:

H1a: Total abnormal trading around a 10-K filing is lower when reports are more complex (e.g., longer/less readable).

Although more complex reports are likely to have some effect on all traders, I predict that the effects of these attributes will be most pronounced in the small investor group. This prediction is consistent with prior experimental research that suggests small (nonprofessional) investors lack investment expertise and have ill-defined valuation models (Maines and McDaniel 2000). Specifically, these non-professional investors tend to read financial statements in the order presented and are therefore more likely to be affected by more complex reports than more sophisticated investors (analysts), who use directed information search strategies to analyze financial statements (Bouwman et al. 1987; Hunton and McEwen 1997).

Although experimental evidence suggests that small investors are most likely to be affected by report attributes, other research indicates that even sophisticated investors are affected by information complexity. For instance, prior experimental research documents that the way financial statement data is presented can influence analysts' judgments (Hopkins 1996; Hirst and Hopkins 1998). Consistent with these experiments, Plumlee (2003) finds that analysts assimilate information in less complex tax changes to greater extent than they assimilate the more complex changes, while a concurrent study by Lehavy et al. (2009) documents that less readable reports are associated with less accurate analyst forecasts and greater dispersion. In sum, these studies suggest that even professional investors may not be immune to report complexity. Thus, although the effects of complex reports are likely to be most pronounced among smaller investors, there is reason to believe that larger, more sophisticated investors could also be affected. Hence, I hypothesize the following (in alternative form):

H1b: The effects of report complexity (i.e., longer/less readable) on abnormal trading are most pronounced among small investors.

Trading Consensus

In addition to trading volume, I also investigate whether more complex reports lead to dispersion in the inferences drawn by those investors who do elect to trade. Specifically, I investigate whether small (large) investors process the data in a different way from other small (large) investors when reports are more complex. This investigation contributes to a large stream of research on the causes and implications of disagreement.[5] Furthermore, this study builds on prior experimental evidence that finds that increases in disclosure lead nonprofessional investors to generate (or infer) private information, while sophisticated professional investors are unaffected by the increased disclosure (Barron et al. 2004). The authors interpret this evidence as larger investors (smaller investors) being more homogeneous (diverse) with respect to their training and

[5] Most analytical models suggest that disagreement of some form drives trading around disclosure releases (Kim and Verrecchia 1991, 1994, 1997; Kandel and Pearson 1995). Karpoff (1986) raises the possibility that this disagreement could stem from information interpreted differently by market agents. Subsequent empirical work by Barron (1995) and Bamber et al. (1997, 1999) provides evidence that the trading around earnings announcements is at least partially attributable to newfound disagreement. Recent analytical work by Bloomfield and Fischer (2009) highlights the importance of studying disagreement, since different forms of disagreement among investors can affect a firm's cost of capital.

occupational selection and therefore more (less) likely to have common interpretations of disclosures. I investigate whether these findings regarding small and large investor behavior hold in an archival market-based study.

As discussed in greater detail later, I measure small and large investors' trading consensus as the absolute value of daily *net buyer initiated* (buyer minus seller initiated) shares traded, deflated by *total shares* (buyer plus seller initiated) traded on the same day. If complexity results in a high level of disagreement within an investor group, then buyer and seller initiated trades would offset and trading consensus would approach 0. Alternatively, if investors agree about the information content of the report, then they are more likely to trade in the same direction and trading consensus would approach 1. Based on these arguments, I hypothesize that, conditional on trade, more complex reports will lead to more disagreement within an investor class. Formally, I hypothesize the following (stated in alternative form):

H2: Trading consensus within a class of investors is lower when 10-K reports are more complex (e.g., longer/less readable).

III. RESEARCH DESIGN

The research design consists of several stages. The first stage examines trading responses to complex 10-K filings using abnormal volume in aggregate as well as the specific trading behavior of small and large investors. The second stage examines the effects of length on small and large investor trading consensus. Finally, I examine whether these associations are driven by cross-sectional variation in firms' disclosure attributes and/or variations in disclosure length over time.

Total Abnormal Trading

To examine the effects of complex reports on the combined trading activity of small, medium, and large investor groups, I estimate the following regression model:

$$\begin{aligned} AVOL_{it} = {} & \alpha_0 + \alpha_1 COMPLEX_{it} + \alpha_2 AFTEAD_{it} + \alpha_3 AFTEXD_{it} + \alpha_4 ABS\_RET_{it} + \alpha_5 PVOL_{it} \\ & + \alpha_6 MV_{it} + \alpha_7 MTB_{it} + \alpha_8 EARNINGS_{it} + \alpha_9 NY\_EARNINGS_{it} + \alpha_{10} NA\_FOLL_{it} \\ & + \alpha_{11} NBSEG_{it} + \alpha_{12} NGSEG_{it} + \varepsilon_{it} \end{aligned} \quad (1)$$

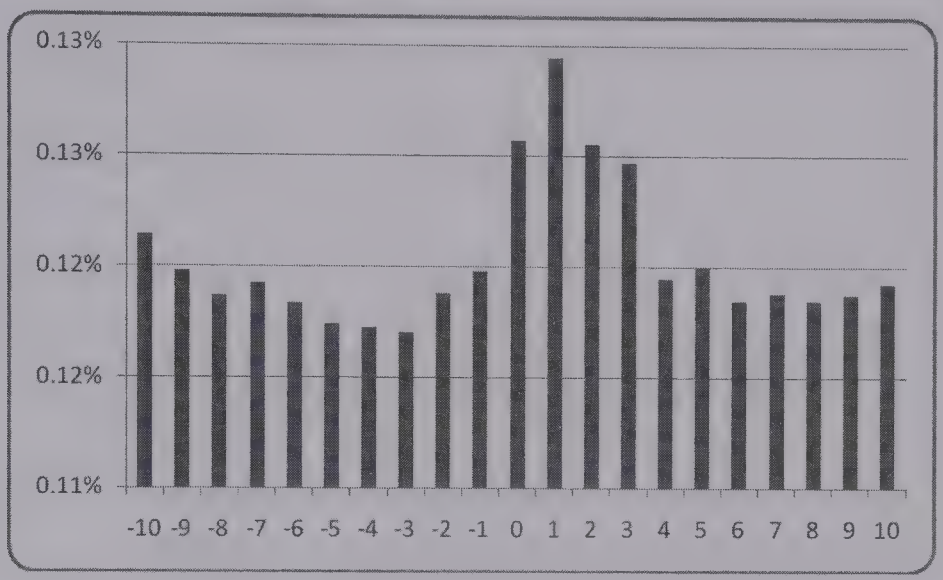
where the dependent variable, *AVOL*, is the measure of abnormal trading volume. The use of abnormal volume (versus raw volume) should mitigate the effect of both technological trading changes (e.g., increased online trading) as well as other factors, such as increases in decimalized trading during the sample period (Barber et al. 2009). Consistent with Asthana et al. (2004), I define abnormal trade, *AVOL*, as the mean daily trading volume during the event period (−1, 3) minus the mean daily trading volume during the non-filing period (−49, −5), deflated by the standard deviation of daily trading volume during the non-filing period (−49, −5).[6]

To determine whether the event windows implemented above are reasonable, I examine the average daily trade volume scaled by shares outstanding classified for each investor group (i.e., small and large traders) in the period surrounding the filing. Figure 1, Panel A provides confirming evidence that most of the small trading volume related to the 10-K filing appears to be captured in

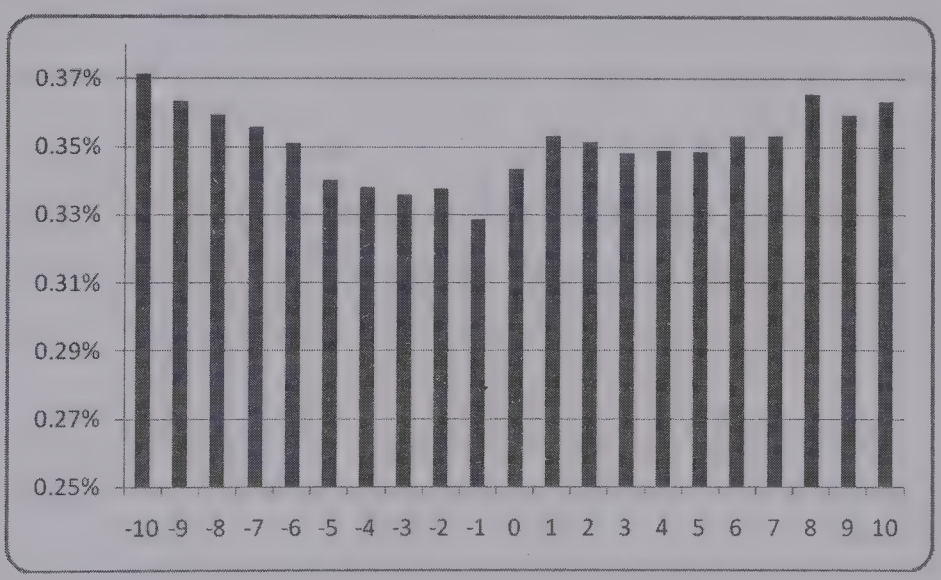
[6] In addition to abnormal volume, I also examine the effects of report complexity on abnormal number of transactions, *ANUM*. Results using the number of transactions are quantitatively similar to the volume results (untabulated). I also measure excess volume as in Barron et al. (2005), where *EXVOL* (*EXNUM*) is defined as the natural log of the cumulative trading volume (transactions) over the five-day event period (−1, 3) minus the natural log of the median volume (transactions) for contiguous five-day periods during the non-filing period (−49, −5). Results using *EXVOL* (*EXNUM*) are consistent with the *AVOL* results, where small investors are significantly affected by more complex reports and (in most cases) small investors are more affected than larger investors.

FIGURE 1
Average Daily Trading Volume around 10-K Filing Dates

Panel A: Small Investors' Daily Trading Volume



Panel B: Large Investors' Daily Trading Volume



Daily trading volume is the number of shares traded scaled by total number of shares outstanding by investor group (i.e., small or large). Panel A (B) plots the mean of small (large) trade volume.

the −1, 3 day event period. Panel B reveals that the trading behavior for large investors appears to be less concentrated around the actual filing date, which is consistent with large investors relying less on the information in the 10-K filing, potentially due to their ability to gain access to

similar information prior to the release of the actual filing. Although the short event window appears to capture the trading specifically related to the 10-K filing, I provide supplementary analysis of longer windows in Section VI.

COMPLEX is the relevant measure of length or readability. Report length is measured as *WORDS* or *WORDS+TABLES*. Because of skewness in the raw number of words, I follow Li (2008) and define *WORDS* as the logarithm of the number of words in the entire document. Although this measure provides a useful metric of the amount of data in text format, it fails to incorporate data included in tables. Therefore, I also examine *WORDS+TABLES*, which is defined as the natural log of the total words plus table cells contained in the document. Since the ability to count the number of table cells requires compliance with a specific HTML standard, the measure is only available for a subset of firms after 1999. Appendix A provides details on the methods used to create these variables.

I first measure readability using the Fog Index, *READ_FOG*, which was originally developed by Gunning (1952). This measure provides a simple, well-known, and widely accepted formula for measuring readability.[7] However, the Fog Index was not developed to specifically measure the attributes of financial reporting complexity. Thus, I also use a proprietary computational software program, *StyleWriter—Plain English Editor*, to develop a unique multidimensional measure of financial readability based on the writing factors outlined by the SEC plain English guidelines, *READ_PE*.[8] As described in more detail in Appendix B, *READ_PE* is designed similar to *READ_FOG* so that higher readability scores are interpreted as less readable reports.

In the first set of tests, I control for potential firm characteristics that may affect abnormal trading. For instance, Asthana et al. (2004) point out the importance of controlling for the effect of the timing of the 10-K filing on trading activity. To control for the effects of earnings information disclosed prior to the 10-K release, I include the variable *AFTEAD*, which is defined as the number of days after the preliminary earnings announcement date that the 10-K is filed. I expect that the longer after the earnings announcement the 10-K is filed, the lower the abnormal trading around the filing.

Prior research also shows that firms delay releasing bad news (Kross and Schroeder 1984) and, consequently, the timeliness of that release affects the market reaction (Chambers and Penman 1984). Based on these findings, I also include *AFTEXD*, which is defined as the number of days the current 10-K is filed after the date of the previous year's 10-K filing. I expect a positive association between the length of the reporting delay after the expected filing date (i.e., previous year's filing date), and abnormal trading volume.

To control for the information content of the report, I include *ABS_RET*, defined as the absolute value of the firm's abnormal (market model) return calculated over the five-day event window $(-1, 3)$. I expect that greater information content will lead to higher abnormal trading volume (Bamber and Cheon 1995; Bamber et al. 1997). I control for uncertainty in the period prior to the 10-K filing by including *PVOL*, defined as the standard deviation of stock price in the period prior to the filing $(-49, -5)$. Theory does not provide a clear prediction on the effects of *PVOL*. I control for firm size by including the independent variable *MV*, the logarithm of the market value

[7] Consistent with prior literature (Li 2008; Collins-Thompson and Callan 2005; Muresan et al. 2006), I measure Fog readability as the ((*words per sentence* + *percent of complex words*) * .4). Additional information on how Fog readability is calculated is provided in Appendix B.

[8] I create the plain English readability measure, defined as the number of errors identified by Stylewriter that match the plain English concerns of the SEC. Specifically, I measure the number of errors defined by the program as [(Passive Verbs + Hidden Verbs + Overwriting + Legal Words & Jargon + Tautologies) * 10] and scale by the approximate number of sentences in the document (number of words/average sentence length). The scalar parses out the effects of the document length, but perhaps more importantly addresses the SEC's concern about sentence length (longer sentences decrease the denominator and thus increase the readability measure).

of equity at the end of the fiscal period. The effects of *MV* on abnormal trading volume are not straightforward. Larger firms are likely to have better information environments (Atiase 1985; Bamber 1987), so firm size could be associated with lower abnormal trading around the filing. However, large firms could also attract more attention-based trading, leading to higher abnormal volume around the filing (Merton 1987). Thus, my prediction on *MV* is nondirectional. I also control for a firm's future growth opportunities on volume, *MTB*, defined as the market value to book value of equity at the close of the current fiscal year. Theory does not provide a clear prediction on *MTB*.

Li (2008) documents that firms with lower earnings tend to have longer and less readable reports. To control for the possibility that *LENGTH* is merely proxying for earnings information, I follow Li (2008) and include *EARNINGS*, which is defined as operating earnings scaled by total assets. Li (2008) also documents that firms with shorter reports tend to have higher earnings persistence. Specifically, Li finds that shorter reports are associated with a greater ability to predict subsequent earnings. In order to address the possibility that report length is simply proxying for information related to next year's earnings, I include *NY_EARNINGS* as a proxy for the expectation of future earnings.[9] Theory does not provide a clear prediction on the effects *EARNINGS* or *NY_EARNINGS* on abnormal volume.

I expect that there will be less new information in the 10-K filings for firms with greater analyst coverage. As such, I include *NA_FOLL*, which I define as the log of 1 plus the number of analysts as reported on I/B/E/S detail and excluded files. My expectation is that, all else equal, higher analyst following will be associated with lower abnormal trade volume. I attempt to control for underlying firm complexity in this analysis by including the logarithm of 1 plus the number of business segments (*NBSEG*) and geographic segments (*NGSEG*), as reported at the end of the fiscal period.[10] Theory does not provide a clear prediction on the effects *NBSEG* or *NGSEG* on abnormal volume. To minimize the effects of outliers, all variables are winsorized by year at the top and bottom 1 percent. Finally, all regressions are performed with clustered robust standard errors (Huber 1967; White 1980) to control for both within firm and within year correlation. Table 1 summarizes all variable definitions.

Small and Large Investors Abnormal Trading

In addition to examining the total trade activity of all investors, I also examine the specific trade activity of small and large investors by estimating the following equation:

$$\begin{aligned} AVOL_{it} = {} & \alpha_0 + SML * [\alpha_1 COMPLEX_{it} + \alpha_2 AFTEAD_{it} + \alpha_3 AFTEXD_{it} + \alpha_4 ABS\_RET_{it} \\ & + \alpha_5 PVOL_{it} + \alpha_6 MV_{it} + \alpha_7 MTB_{it} + \alpha_8 EARNINGS_{it} + \alpha_9 NY\_EARNINGS_{it} \\ & + \alpha_{10} NA\_FOLL_{it} + \alpha_{11} NBSEG_{it} + \alpha_{12} NGSEG_{it}] + \beta_0 + LRG * [\beta_1 COMPLEX_{it} \\ & + \beta_2 AFTEAD_{it} + \beta_3 AFTEXD_{it} + \beta_4 ABS\_RET_{it} + \beta_5 PVOL_{it} + \beta_6 MV_{it} + \beta_7 MTB_{it} \\ & + \beta_8 EARNINGS_{it} + \beta_9 NY\_EARNINGS_{it} + \beta_{10} NA\_FOLL_{it} + \beta_{11} NBSEG_{it} \\ & + \beta_{12} NGSEG_{it}] + \varepsilon_{it}. \end{aligned} \quad (2)$$

[9] Perfect foreknowledge of next year's earnings is admittedly a strong assumption. However, the use of next year's earnings as a proxy for persistence (ability of report attributes to predict next year's earnings) does not appear to bias toward finding that more complex reports affect trading behavior. Further, the exclusion of this control variable does not affect the results.

[10] In addition to *NBSEG* and *NGSEG*, I also add controls to all regressions for whether a firm had a merger during the year as this would also increase firm complexity. I find that the merger control does not affect the significance of the results reported in this study (untabulated).

TABLE 1
Variable Definitions

Variable	Definitions (Alphabetical Order)
ABS_RET	Market adjusted return, defined as the absolute value of the abnormal (market model) share price return over the event period (−1, 3).
AFTEAD	Number of days after the earnings announcement date, defined as the number of calendar days the 10-K is filed after the preliminary earnings announcement date.
AFTEXD	Number of days after the expected filing date, defined as the number of calendar days the 10-K is filed after the expected filing date (last year's 10-K filing date).
AVOL	Abnormal trading volume, defined as the mean daily trading volume during the event period (−1, 3) minus the mean daily trading volume during the non-filing period (−49,−5), deflated by the standard deviation of daily trading volume during the non-filing period (−49,−5).
COMPLEX	Relevant measure of reporting complexity (i.e., *WORDS*, *WORDS+TABLES*, *READ_FOG*, or *READ_PE*).
EARNINGS	Operating earnings scaled by assets, defined as Compustat items (data178/data6).
LRG	Large trades: When suffixed to a variable, implies large trades (greater than or equal to $50,000).
NA_FOLL	Number of analysts following the firm, natural logarithm of 1 plus the number of unique analysts providing a forecast from the I/B/E/S Detail and Excluded Files.
MV	Market value of equity, defined as the natural logarithm of market value of common equity at the close of the current fiscal year (data25 * data199) from Compustat.
MTB	Market-to-book, defined as the market value of the firm's equity divided by its book value ((data24 * data25)/data60) from Compustat.
NBSEG	Number of business segments, defined as the natural log of 1 plus the number of business segments.
NGSEG	Number of geographic segments, defined as the natural log of 1 plus the number of geographic segments.
NY_EARNINGS	Next year's operating earnings scaled by assets, defined as Compustat items (data 178/data 6).
PVOL	Price volatility, defined as the standard deviation of the firm's stock price prior to the 10-K filing (−49, −5).
READ_FOG	Fog index, calculated as [(words per sentence + percent of complex words) * .40] using Perl's En Fathom module.
READ_PE	Plain English index, calculated using Stylewriter software as described in Appendix 2 as 10 * ((# of errors attributable to: Passive Verbs + Hidden Verbs + Overwriting + Legal Words & Jargon + Tautologies)/(number of sentences)).
SML	Small trades: When suffixed to a variable, implies small trades (if share price is < $50, then dollar amount of trade is less than or equal to $5,000; if share price >50, but less than or equal to 100, then dollar amount of trade is less than or equal to 100 * price; if share price is greater than 100, then the firm is excluded from the sample).
WORDS+TABLES	Natural logarithm of the total of number of words plus the number of table cells included in each document.

(continued on next page)

TABLE 1 (continued)

Variable	Definitions (Alphabetical Order)
TRADE_ CONS	Trade Consensus, defined as the absolute value of the daily net buys during the event period (buys minus sales), deflated by the total number of shares traded (buys plus sales) during the event period.
WORDS	Natural logarithm of the number of words included in the 10-K report.

Consistent with Bhattacharya (2001), trades are assumed to be made by small investors if the dollar amount of the trade is less than or equal to $5,000; when the dollar amount of the trade is greater than or equal to $50,000, the trade is assumed to be made by a large investor. Consistent with prior literature, I ignore medium-sized trades to create a "buffer-zone" between the trading activities of small and large investors.[11] Consistent with Asthana et al. (2004), I modify this formulation so that for firms with a share price greater than $50, I define small trades as those less than or equal to 100 times the share price.[12] However, in order to keep a sufficient buffer between small and large trades, I exclude firms with a share price greater than $100 (i.e., the maximum small trade in my sample is $10,000).

Equation (2) represents the stacking of two regressions that allows statistical tests of differences in coefficient estimates across the small and large investor groups. *SML* (*LRG*) is an indicator variable equal to 1 when the dependent variable represents the abnormal trading behavior for small (large) investors, and 0 otherwise. Accordingly, the α (β) coefficients measure associations between small (large) investor trade behavior and the independent variables.[13] My expectation for both small and large trade groups is that the coefficients on *COMPLEX* will be negative if increases in complexity result in increased processing costs leading to lower abnormal trade. I also examine whether small investors are more likely to be affected by more complex reports. Accordingly, I provide a test of coefficients across *SML* * *COMPLEX* and *LRG* * *COMPLEX* and expect the effects of report complexity to be greater for small investors.

Trading Consensus

I also examine whether increases in report complexity lead investors within the same class (e.g., small investors) to draw different inferences from the same reported information. To gauge the amount of trade consensus by investor group around a 10-K filing, I first classify all buyer and seller initiated trades using the standard Lee and Ready (1991) algorithm.[14] Next, I create a

[11] Lee and Radhakrishna (2000) find that these trade size bins perform well in identifying trades initiated by individual investors and institutions. Eliminating medium-sized trades increases the power of the test, since large investors may try to break up their trades to disguise their identity (Kyle 1985; Meulbroek 1992; Barclay and Warner 1993) but for a variety of reasons are unlikely to make very small trades (Bhattacharya et al. 2007).

[12] Consistent with prior literature, I also exclude the opening trade because it is often the sum of multiple orders and including it could add noise to the measures (Lee and Ready 1991; Lee 1992; Bhattacharya et al. 2007). Further, I only include trades with a "regular sales" condition code. Bhattacharya et al. (2007, 587) point out that these "trades result from continuous two-sided auctions involving market orders, limit orders, and buys and sells against the specialists' inventories. This is not the case when the condition code indicates something other than a 'regular sale' (e.g., large block trades or stopped orders)."

[13] See Maddala (2001) for discussion of stacked regressions. Since the technique captures any (potential) correlations across the error terms, stacking allows statistical tests of coefficients across equations (i.e., small and large investor groups).

[14] This algorithm, validated by Lee and Radhakrishna (2000), classifies trades above (below) the bid/ask midpoint as buyer (seller) initiated. Trades occurring at the midpoint are classified using the tick test, which compares the trade price to

variable *TRADE_CONS*, defined as the absolute value of daily net buyer initiated shares traded (buyer minus seller initiated) deflated by total volume of shares traded (absolute value of buyer plus seller initiated) on the same day for each investor group (e.g., small).[15] I then replace the dependent variable, *AVOL*, in Equation (2) with *TRADE_CONS*.

If small or large investors process the 10-K and agree (disagree) on the information's effect on the firm's terminal value, they are likely to trade in a similar (different) direction to other investors in that same group leading to higher (lower) *TRADE_CONS*. As predicted in H2, I expect more complex reports to increase disagreement for both small and large investor groups. Hence, I predict negative coefficients on all measures of *COMPLEX*.

IV. SAMPLE AND DESCRIPTIVE DATA

Sample Collection

I gather the sample by downloading all 10-K and 10-K405 filings on EDGAR from 1994 to 2006.[16] I require a firm-year match on PERMNO (from CRSP), GVKEY (from Compustat), CIK (from EDGAR), and TAQ data for small and large volume around the 10-K filing.[17] Consistent with prior research, I eliminate observations where the ticker identifier on TAQ changed during the year, the stock price was less than one dollar, or the firm had a stock split or issued a stock dividend during the year. Using the process described in Appendix A, I also eliminate all observations with insufficient text remaining after I eliminated the header, table, and other information. In order to calculate the *AFTEXD* variable, I require the prior year filing date to be available on EDGAR, which eliminates all 1994 observations and a significant number of other observations. Finally, I eliminate observations where there is missing Compustat data (e.g., *MV*, *ASSETS*) or where analyst following is unavailable on I/B/E/S. The final sample consists of 3,809 unique firms with 12,771 firm year observations. This number reduces to 4,724 observations for *WORDS +TABLES* analysis. Table 2 provides a detailed description of the sample selection procedure.

Sample Characteristics

Table 3 provides sample characteristics. Panel A provides descriptive statistics for several variables of interest. The positive aggregate mean for *AVOL* is 0.17, and the means for small investor *AVOL* and large investor *AVOL* are 0.22 and 0.18, respectively. The positive value for small investor *AVOL* is consistent with survey evidence small traders appear to trade on the information contained in the 10-K.[18] The mean for small investor *TRADE_CONS* is 0.24, whereas the mean for large investor *TRADE_CONS* is 0.52. These averages are consistent with the general notion that small traders are more likely to have disagreement surrounding information events than

---

adjacent trades. Under the tick test, a trade is classified as buyer (seller) initiated when the price is higher (lower) than the price of the previous trade. In cases when the price is the same as the previous trade (a zero tick), the classification of buyer (seller) initiated trades defaults to the last trade where there was a price change.

[15] The limit order trader on the other side of the buyer or seller initiated trade could be an investor from the same or a different size class. For example, the other side of a small buyer or seller initiated trade could be another small investor, the market maker, or a large or medium passive trader whose limit order is broken up into smaller lots. The important factor is not whether the limit order investor on the other side of the buyer/seller initiated transaction is a small or large trader, but rather, whether the transaction was initiated as a buy or sale by a small or large trader.

[16] Li and Ramesh (2009) show that the stock market and volume reaction for quarterly and small business filings disappear after controlling for the concurrent release of earnings information. As such, I exclude small business filers and restrict my investigation to annual filings. Forms 10-K and 10-K405 are identical in substance, except that form 10-K405 indicates that an officer or director failed to file a Form 4 (or Form 3 or 5) within the required time period.

[17] For comparative purposes across small and large investors, I ensure that each firm analyzed has both small and large trade activity by eliminating a trivial number of observations that have no trade for an investor type during the control and five-day window around the 10-K filing. The results are unaffected by the inclusion of these observations.

[18] All *AVOL* coefficients are negative at the median, which is consistent with Asthana et al. (2004) and consistent with the earnings releases having a significant impact on pre-period volume.

TABLE 2

Sample Selection

	No. Firm-Year Observations
10-K Filings from EDGAR (1994–2006) with GVKEY and PERMO Identifiers and TAQ data	21,189
less where the TAQ ticker changed during the year	(129)
less where there was a stock split or stock dividend	(1,523)
less where the stock price was less than $1	(122)
less where the filing is missing *LENGTH* and FOG from Perl En Fathom	(4)
less where there is missing prior year filing date (*AFTEXD*)	(3,763)
less missing Compustat data (*MV*, *ASSETS*, *NBSEG*, *NGSEG*, etc.)	(1,002)
less missing I/B/E/S data to calculate analyst following	(1,875)
Observations available for primary analysis	12,771
Observations available for # of table cells in 10-K available (*WORDS+TABLES*) calculation	4,724

The ability to count the number of table cells requires firms to comply with a specific HTML standard. Hence, *WORDS+TABLES* is only available for a subset of firms after 1999.

larger more sophisticated investors. The mean *ABS_RET* is 4.8 percent, indicating an abnormal market reaction during the 10-K filing period. The mean *WORDS* is 10.39, which translates into an average of 32,533 words per 10-K, while the mean Fog Index is 19.94, which classifies the vast majority of financial statements as unreadable according to most interpretations of the index.

Panel B (C) of Table 3 provides information on report length (readability) across time. Consistent with Li (2008), there is a significant increase in *WORDS* and *WORDS+TABLES* over the sample period investigated.[19] Consistent with the increase in length, Panel C reveals that there is also a significant increase in *READ_FOG*, indicating that reports have become less readable over time. However, this pattern in readability is not evident in *READ_PE*, where there is no significant difference in readability during the sample period. In summary, there appears to be a significant increase in report length over time, but the evidence for a corresponding change in readability is only evident when the Fog measure of readability is employed.

V. REGRESSION RESULTS

Total Trading Activity

Table 4 reports the multivariate regression results for the effects of complexity on abnormal trading volume (*AVOL*). Overall, I find evidence consistent with more complex reports affecting total trading volume. Specifically, both coefficients related to longer reports as reported in columns (1) and (2) of Panel A in Table 4 are significantly negative ($p < .01$).[20] While both measure of readability are negative, only the readability measure based on plain English attributes as prescribed by the SEC is significant ($p < .01$).

Although the majority of control variables are significant in the expected direction in Table 4, Panel A, some of the controls warrant further discussion. Specifically, the significant positive

[19] In untabulated analysis, I find a high correlation between *WORDS* and *WORDS+TABLES* (85 percent Pearson and 96 percent Spearman), which suggests that the quantity of data in tables is proportional to the length of text.

[20] All p-values reported in the study are presented as two-tailed when no prediction is given and one-tailed when predicted.

TABLE 3

Sample Characteristics

Panel A: Distribution of Variables

Variables	# Obs.	Mean	Std. Dev.	25th Percentile	Median	75th Percentile
AVOL	12771	0.173***	1.013	−0.402	−0.104#	0.403
SML**AVOL*	12771	0.222***	1.030	−0.412	−0.028###	0.573
LRG**AVOL*	12771	0.177***	1.026	−0.366	−0.121###	0.338
SML**TRADE_CONS*	12771	0.244***	0.167	0.125	0.198###	0.315
LRG**TRADE_CONS*	12771	0.521***	0.292	0.272	0.472###	0.750
WORDS	12771	10.391***	0.790	9.967	10.384###	10.845
WORDS+TABLES	4724	10.828***	0.569	10.508	10.816###	11.164
READ_FOG	12771	19.943***	1.782	18.842	19.762###	20.869
READ_PE	12771	21.157***	5.240	17.546	20.002###	23.708
AFTEAD	12771	38.431***	17.927	28.000	41.000###	51.000
AFTEXD	12771	−1.293***	10.322	−5.000	0.000###	3.000
ABS RET	12771	0.048***	0.053	0.014	0.031###	0.063
PVOL	12771	1.810***	2.137	0.675	1.203###	2.106
MV ($ billion)	12771	2.752***	7.445	0.232	0.605###	1.791
MTB	12771	1.603***	1.709	0.581	1.062###	1.938
EARNINGS	12771	0.052***	0.155	0.030	0.077###	0.125
NY_EARNINGS	12771	0.042***	0.208	0.025	0.073###	0.121
NBSEG	12771	0.693***	0.752	0.000	0.000###	1.386
NGSEG	12771	1.012***	0.631	1.099	1.099###	1.386
NA_FOLL	12771	2.065***	0.758	1.609	2.079###	2.639

(continued on next page)

Panel B: Mean and Median Length across Time

	# Obs.		*WORDS*		*WORDS+TABLES*	
Year	***WORDS***	***WORDS+TABLES***	**Mean**	**Median**	**Mean**	**Median**
1995	312	—	10.147	10.091	—	—
1996	437	—	10.136	10.088	—	—
1997	905	—	10.164	10.115	—	—
1998	1514	—	10.234	10.146	—	—
1999	1393	6	10.315	10.234	10.496	10.302
2000	1121	72	10.319	10.231	10.484	10.432
2001	1031	175	10.332	10.271	10.446	10.467
2002	1151	495	10.295	10.394	10.592	10.625
2003	1124	751	10.536	10.520	10.814	10.796
2004	1386	1098	10.601	10.531	10.878	10.825
2005	1353	1170	10.612	10.595	10.908	10.885
2006	1044	957	10.611	10.600	10.904	10.896
Change in *WORDS* 2006 less 1995			0.464[a]	0.509[b]		
Change in *WORDS+TABLES* 2006 less 2000					0.408[a]	0.594[b]

(continued on next page)

Panel C: Mean and Median Readability across Time

Year	# OBS	*READ_FOG*		*READ_PE*	
		Mean	Median	Mean	Median
1995	312	19.603	19.303	21.126	20.421
1996	437	19.660	19.462	21.362	20.091
1997	905	19.681	19.517	21.388	20.501
1998	1514	19.848	19.729	21.594	20.377
1999	1393	19.759	19.560	21.329	20.158
2000	1121	19.684	19.497	21.196	19.769
2001	1031	19.561	19.351	20.757	19.274
2002	1151	19.583	19.551	20.789	19.659
2003	1124	20.149	19.918	21.241	20.125
2004	1386	20.465	20.151	21.314	20.133
2005	1353	20.380	20.311	20.905	19.912
2006	1044	20.338	20.100	20.795	19.781
Change in FOG Readability 2006 less 1995		0.735[c]	0.509[d]		
Change in Plain English Readability 2006 less 1995				−0.331	−0.640

*, **, *** (#, ##, ###) Indicate two-tailed statistical significance of whether the mean (median) coefficient estimates differ from 0 at the 10 percent, 5 percent, and 1 percent levels, respectively, based on t-tests (Wilcoxon signed rank tests).

[a,b] Represents significant differences at the 1 percent level in *LENGTH* between the corresponding years tested using t-test (Wilcoxon rank tests). Because of the minimal number of observations for *WORDS+TABLES* in 1999, the analysis of changes in *WORDS+TABLES* over time uses 2000 as the base year.

[c,d] Represents significant differences at the 1 percent level in FOG Readability between the corresponding years tested using t-test (Wilcoxon rank tests).

See Table 1 for variable definitions.

TABLE 4

The Effects of Reporting Complexity on Total Abnormal Volume

Panel A: Pooled Analysis

	Hyp. Sign	*WORDS* (1)	*WORDS+TABLES* (2)	*READ_FOG* (3)	*READ_PE* (4)
COMPLEX	−	−0.0284***	−0.0754***	−0.0008	−0.0037***
		(−2.60)	(−3.56)	(−0.15)	(−2.32)
AFTEAD	−	−0.0073***	−0.0113***	−0.0071***	−0.0072***
		(−4.37)	(−5.17)	(−4.46)	(−4.43)
AFTEXD	+	0.0022*	0.0046***	0.0022*	0.0022*
		(1.32)	(2.33)	(1.32)	(1.33)
ABS_RET	+	6.2528***	9.7724***	6.2465***	6.2494***
		(5.88)	(5.83)	(5.88)	(5.87)
PVOL	?	−0.0022	−0.0362**	−0.0020	−0.0020
		(−0.23)	(−1.88)	(−0.21)	(−0.21)
MV	+/−	0.0377***	0.0483***	0.0347***	0.0345***
		(3.88)	(2.44)	(3.71)	(3.65)
MTB	?	−0.0255***	−0.0272**	−0.0240***	−0.0249***
		(−3.58)	(−1.81)	(−3.47)	(−3.45)
EARNINGS	?	0.0445	−0.1166	0.0660	0.0625
		(0.36)	(−0.47)	(0.53)	(0.50)
NY_EARNINGS	?	0.1056	0.4194**	0.1070	0.1052
		(1.34)	(2.32)	(1.34)	(1.32)
NA_FOLL	−	−0.1027***	−0.0923***	−0.1047***	−0.1039***
		(−5.34)	(−2.79)	(−5.36)	(−5.32)
NBSEG	?	0.0119	0.0384***	0.0105	0.0104
		(0.59)	(2.51)	(0.53)	(0.53)
NGSEG	?	−0.0201	−0.0445	−0.0194	−0.0200
		(−1.04)	(−1.34)	(−0.99)	(−1.04)
R^2		12.20%	18.47%	12.16%	12.20%

(continued on next page)

Panel B: Within-Year Analysis

	Hyp. Sign	*WORDS* (1)	*WORDS+TABLES* (2)	*READ_FOG* (3)	*READ_PE* (4)
COMPLEX	−	−0.0302***	−0.0878***	−0.0023	−0.0035***
		(−2.33)	(−2.94)	(−0.46)	(−2.22)
CONTROLS		Yes	Yes	Yes	Yes
R^2		12.71%	18.88%	12.66%	12.64%

Panel C: Within-Firm Analysis

	Hyp. Sign	*WORDS* (1)	*WORDS+TABLES* (2)	*READ_FOG* (3)	*READ_PE* (4)
COMPLEX	−	−0.0203**	−0.0679**	0.0007	−0.0027*
		(−1.75)	(−2.24)	(0.10)	(−1.49)
CONTROLS		Yes	Yes	Yes	Yes
R^2		11.06%	16.69%	11.04%	11.05%

*, **, *** Indicate two-tailed statistical significance of coefficient estimates at the 10 percent, 5 percent, and 1 percent levels, respectively, when no prediction is given and one-tailed significance when predicted.

This table reports the regression results for total (i.e., small, medium, and large investors combined) abnormal trading volume (*AVOL*).

Panel A reports results from estimating Equation (1). Panel B reports results from estimating Equation (1) after adding in year fixed effects. Panel C reports results from estimating Equation (1) after adding in firm fixed effects.

Columns 1, 2, 3, 4 report results when *COMPLEX* is measured as *WORDS*, *WORDS+TABLES*, *READ_FOG*, *READ_PE*, respectively. There is one observation per firm year for a total of 12,771 for all regressions except in column 2 where there are 4,724 observations when *WORDS+TABLES* is the variable of interest. All regressions include a constant term. Regressions are performed with clustered robust standard errors (Rogers 1993) to control for within-firm and -year correlation, except when a specific fixed effect is already included in the model.

coefficient on *MV* suggests that firm size is likely to be correlated with more familiarity and, thus, greater abnormal trading surrounding the filing. The coefficient on *MTB* is significantly negative in all regressions, suggesting that 10-K filings may contain less relative information to traders investing in growth firms. The remaining control variables appear to be significant in the predicted direction or do not exhibit enough of a consistent pattern to warrant discussion.

The analysis in Table 4, Panel A documents an association between report complexity and abnormal trading volume. However, it is unclear whether this association is driven by cross-sectional variation in firms' disclosure attributes and/or variations in disclosure complexity over time. Figure 2 reveals that longer and less readable reports tend to be concentrated in industries that are generally more complex (e.g., utilities or financial institutions). To address whether cross-sectional differences in length and readability lead to abnormal timing while holding time constant, I add year fixed effects to regression 1. The results in Table 4, Panel B are consistent with those in Panel A, where *LENGTH*, *WORDS+TABLES*, and *READ_PE* are all significantly negative ($p < .01$). It appears that the association between report complexity and abnormal trading is attributable to cross-sectional differences in firms' reporting after controlling for time.

Based on the previous analysis, it appears that firms with more complex reports are associated with less trading volume. However, the question of whether the significant increase in report length during the sample period affects investor trading behavior remains. In order to examine whether variations in reporting complexity over time impact trading volume after holding cross sectional differences in disclosure constant, I add firm fixed effects to regression Equation (1).[21] This analysis enables me to investigate whether the variations of a particular firm's report length over time lead to variations in trading volume and transactions for that particular firm's investors. Table 4, Panel C provides the results from the within-firm analysis, where the coefficients on *WORDS*, *WORDS+TABLES*, and *READ_PE* are once again significantly negative (although slightly weaker than in other panels).

Overall, the results in Table 4 provide evidence consistent with a decrease in total trading activity when reports are more complex.[22] Within-year and within-firm analyses suggest that the association between more complex reports and lower abnormal trading is driven by both cross sectional differences in firm reporting complexity as well as variation in disclosure complexity over time. From an economic perspective, after controlling for firm fixed effects, I find that a 55-page increase in a firms 10-K filing (roughly equivalent to increase in report length from 1995 to 2006) results in a 10.1 percent reduction in total abnormal volume around the filing.[23]

Small and Large Investor Trading Activity

Table 5 reports the multivariate regression results for the effects of report complexity on small and large investors' abnormal trading volume. The evidence is consistent with the results from Table 4 being driven by small investor trade. Columns (1) and (2) of Panel A show that the effects

---

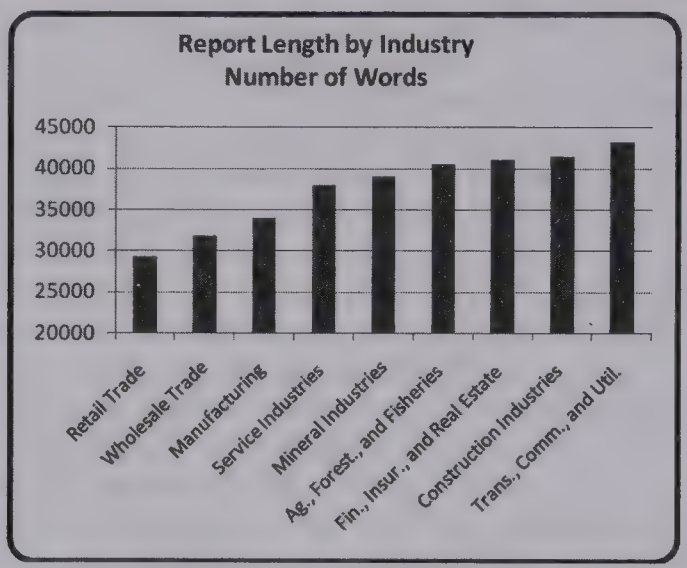
[21] In my primary analyses, I examine firm and year fixed effects separately, as including both effects simultaneously would likely eliminate any meaningful variation in the complexity measures. Specifically, including both effects would provide a test of the effects of report complexity on trading activity relative to the average complexity of the same firm and relative to the average of other firms in the same year. Consistent with controlling away firms' increases in complexity measures over time, the results after including both fixed effects in the same model are greatly diminished. While the effects of longer reports on total and small investor abnormal volume remaining significant at conventional levels, all the analyses involving readability and trading consensus are insignificant.

[22] In untabulated results, I examine the effects of complexity on the short-term abnormal market reaction and find significant evidence that longer reports (and to a lesser extent less readable reports) have more muted market reactions. This is consistent with Bloomfield's (2002) and You and Zhang's (2009) long-window return evidence that information that is hard to extract from financial statements will not be reflected immediately in stock prices.

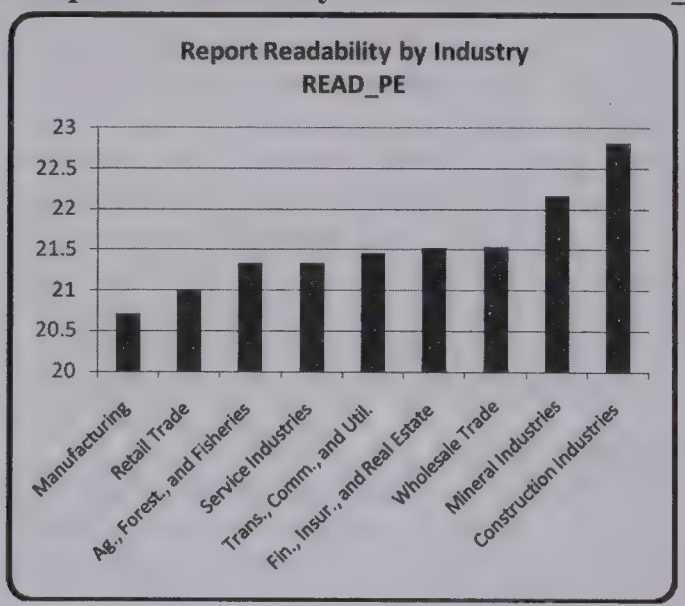
[23] The number of words increased from a median of approximately 24,000 words in 1995 to a median of approximately 39,000 words in 2006. The page estimate provided assumes 270 words per page, which is the approximate number of words contained on a standard 8-1/2 by 11 page that is double-spaced with 12-point font.

FIGURE 2
Report Length by Industry Classification

Panel A: Report Length Measured as Number of Words



Panel B: Report Readability Measured as READ_PE (Plain English)



This figure provides a summary of average report length by the major SIC code groups as identified by the U.S. Census Bureau (http://www.census.gov/epcd/naics/nsic2ndx.htm). Panel A (B) plots the average report length using number of words (plain English readability).

of longer reports are significant for small investors, wherein both the coefficients on *SML* * *WORDS* and *SML* * *WORDS+TABLES* are significant ($p < .01$). There is marginally significant evidence that large investors are affected by longer reports, reported in column (2) ($p = .09$).

TABLE 5

The Effects of Reporting Complexity on Small and Large Investor Abnormal Volume

Panel A: Pooled Analysis

	Hyp. Sign	*WORDS*		*WORDS+TABLES*		*READ_FOG*		*READ_PE*	
		SMALL	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*
		(1)		(2)		(3)		(4)	
COMPLEX	−	−0.0492***	0.0114	−0.0939***	−0.0333*	−0.0133***	0.0153	−0.0092***	0.0026
		(−4.00)	(1.17)	(−3.79)	(−1.37)	(−2.72)	(3.18)	(−4.07)	(1.29)
AFTEAD	−	−0.0062***	−0.0057***	−0.0099***	−0.0086***	−0.0061***	−0.0055***	−0.0067***	−0.0051***
		(−3.39)	(−5.63)	(−3.53)	(−7.16)	(−3.39)	(−5.46)	(−3.89)	(−4.85)
AFTEXD	+	0.0023***	0.0013	0.0026***	0.0033*	0.0023***	0.0012	0.0026***	0.0010
		(2.46)	(0.78)	(2.04)	(1.35)	(2.53)	(0.73)	(3.02)	(0.61)
ABS_RET	+	7.1042***	4.3407***	8.9660***	6.9747***	7.0712***	4.3668***	6.8853***	4.5562***
		(7.46)	(5.65)	(6.62)	(6.16)	(7.51)	(5.70)	(7.43)	(5.70)
PVOL	?	0.0062	−0.0041	−0.0042	−0.0389**	0.0063	−0.0038	0.0076	−0.0051
		(1.21)	(−0.44)	(−0.24)	(−2.39)	(1.24)	(−0.42)	(1.47)	(−0.54)
MV	+/−	0.0776***	−0.0054	0.0849***	0.0002	0.0718***	0.0038	0.0507***	0.0173
		(5.06)	(−0.63)	(3.36)	(−0.01)	(4.86)	(−0.41)	(3.36)	(1.52)
MTB	?	−0.0179***	−0.0160***	−0.0364***	−0.0168*	−0.0156**	−0.0162***	−0.0182**	−0.0153**
		(−2.68)	(−2.51)	(−4.08)	(−1.68)	(−2.31)	(−2.58)	(−2.51)	(−2.43)
EARNINGS	?	0.3071***	−0.0763	0.0142	0.0113	0.3383***	−0.0769	0.3516***	−0.0986
		(3.84)	(−0.76)	(0.11)	(0.06)	(4.53)	(−0.76)	(4.61)	(−0.97)
NY_EARNINGS	?	−0.1346**	0.1826***	0.1858	0.2497*	−0.1333**	0.1837***	−0.1230**	0.1694***
		(−2.33)	(4.60)	(1.61)	(1.91)	(−2.22)	(4.57)	(−2.03)	(4.27)
NA_FOLL	−	−0.0813***	−0.0885***	−0.0614	−0.0827***	−0.0836***	−0.0891***	−0.0752***	−0.0959***
		(−2.72)	(−5.12)	(−1.10)	(−2.71)	(−2.77)	(−5.09)	(−2.49)	(−5.40)
NBSEG	?	0.0175	0.0195	0.0387***	0.0437*	0.0146	0.0205	0.0161	0.0187
		(0.76)	(1.14)	(2.85)	(1.83)	(0.64)	(1.21)	(0.72)	(1.11)
NGSEG	?	−0.0208	−0.0078	−0.0231	−0.0335	−0.0219	−0.0054	−0.0287	0.0000
		(−0.88)	(−0.51)	(−1.09)	(−1.15)	(−0.92)	(−0.34)	(−1.16)	(0.00)
R^2			10.36%		13.84%		10.33%		10.16%
Test of (*SML* * *COMPLEX* − *LRG* * *COMPLEX*) < 0			−0.061###		−0.061###		−0.029###		−0.012###

(continued on next page)

Panel B: Within-Year Analysis

	Hyp. Sign	*WORDS*		*TABLES + WORDS*		*READ_FOG*		*READ_PE*	
		SMALL	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*
		(1)		(2)		(3)		(4)	
COMPLEX	−	−0.0443***	0.0054	−0.0648***	−0.0657**	−0.0093**	0.0120	−0.0085***	0.0027
		(−4.43)	(0.56)	(−2.49)	(−2.66)	(−2.26)	(2.91)	(−5.69)	(1.87)
CONTROLS			Yes		Yes		Yes		Yes
R^2			11.01%		14.34%		10.97%		10.71%
Test of (*SML* * *COMPLEX* − *LRG* * *COMPLEX*) < 0			−0.050###		0.001		−0.021###		−0.011###

Panel C: Within-Firm Analysis

	Hyp. Sign	*WORDS*		*WORDS+TABLES*		*READ_FOG*		*READ_PE*	
		SMALL	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*
		(1)		(2)		(3)		(4)	
COMPLEX	−	−0.0269**	0.0029	−0.1018***	−0.0501*	0.0002	0.0044	−0.0026*	−0.0023
		(−2.06)	(0.29)	(−3.01)	(−1.49)	(0.02)	(0.77)	(−1.62)	(−1.23)
CONTROLS			Yes		Yes		Yes		Yes
R^2			9.41%		12.58%		9.40%		9.10%
Test of (*SML* * *COMPLEX*−*LRG* * *COMPLEX*) < 0			−0.030##		−0.052		−0.004		0.000

*, **, *** Indicate two-tailed statistical significance of coefficient estimates at the 10 percent, 5 percent, and 1 percent levels, respectively, when no prediction is given and one-tailed significance when predicted.

#, ##, ### Indicate one-tailed statistical significance of differences between small and large investor coefficients for *COMPLEX* using F-statistics.

This table reports the regression results for small and large investor abnormal trading volume (*AVOL*).

Panel A reports results from estimating Equation (2). Panel B reports results from estimating Equation (2) after adding in year fixed effects. Panel C reports results from estimating Equation (2) after adding in firm fixed effects..

Columns 1, 2, 3, 4 report results when *COMPLEX* is measured as *WORDS*, *WORDS+TABLES*, *READ_FOG*, *READ_PE*, respectively. All regressions include a constant term. The reported regressions are stacked regressions of small and large investor groups, where there are two observations (i.e., small and large) for each firm year. Hence, there are a total of 25,542 for each regression, except in column 2 where there are 9,448 observations when *WORDS+TABLES* is the variable of interest. To minimize the effects of outliers, all variables are winsorized at the top and bottom 1 percent by year. Regressions are performed with clustered robust standard errors (Rogers 1993) to control for within-firm and -year correlation, except when a specific fixed effect is already included in the model.

Further, small investors are more affected by longer reports than large investors as the differences in coefficients in columns (1) and (2) on *SML* * *COMPLEX* and *LRG* * *COMPLEX* are both significant ($p < .01$).

I find that the effects of less readable reports also appear to be driven by small investors. Specifically, columns (3) and (4) of Table 5, Panel A show the effects of less readable reports for small investors, where both the coefficients on *SML* * *READ_FOG* and *SML* * *READ_PE* are both significant ($p < .01$). There is no evidence of less readable reports reducing the trading behavior of large investors. Further, small investors are more affected by less readable reports than large investors; the differences in coefficients in columns (3) and (4) on *SML* * *COMPLEX* and *LRG* * *COMPLEX* are both significant ($p < .01$).

As with Table 4, most of the control variables in Table 5 are significant in the expected direction. However, some of the specific coefficients on small and large investors warrant additional discussion. Specifically, coefficients on *MV* are significantly positive only for small investors, which is consistent with small investors trading more around 10-K filings that are larger potentially because these firms are more familiar. The coefficients on *MTB* are significantly negative for both small and large investors, suggesting that the 10-K filing for growth firms contains less pertinent information for both sets of investors. The coefficients on *EARNINGS* are positive for small trade activity in three of the four regressions, which indicates that small investors trade more when earnings are higher. Furthermore, the coefficients on *NY_EARNINGS* are significantly negative (positive) for small (large) investors in most regressions. These differences suggest that small (large) investors' trade less (more) when the filings contain information related to positive future earnings, which could be interpreted as more sophisticated investors taking more of a long-term perspective.

Another interesting insight provided in Table 5 are the coefficients on *LRG* * *COMPLEX* when *READ_FOG* and *READ_PE* are investigated. The positive association between the readability measures and increased abnormal trading is inconsistent with the prediction and appears to be consistent with large investors being attracted to more complex reports. This attraction could be due to large investors trading more either because they have a processing advantage over other investors, or because certain firms cater to institutional investors by making their reports more technical.

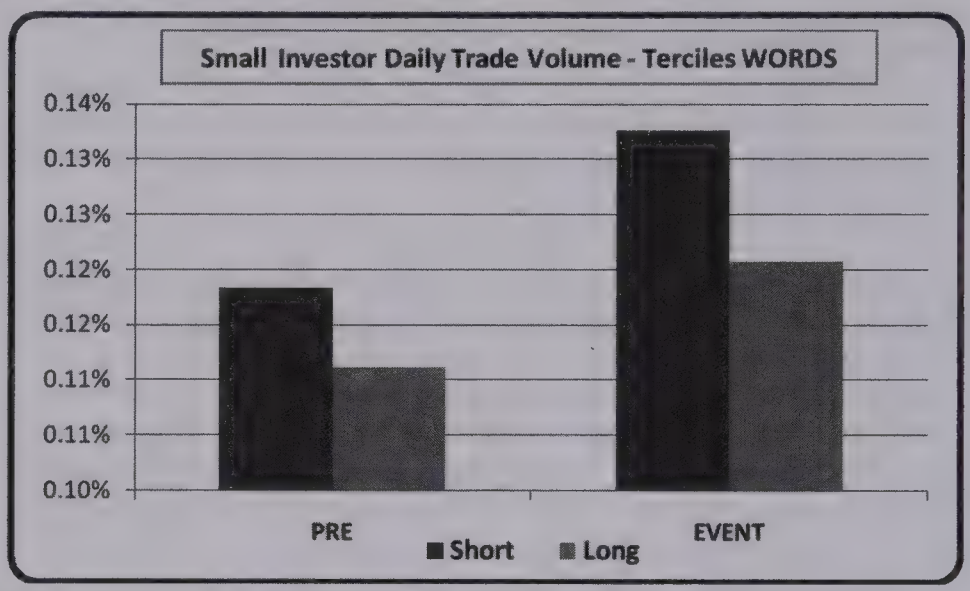
To address whether this association is consistent with investors' overall trading strategies, I investigate whether large (small) investors are associated with firms that have more (less) complex filings. I first create a measure of trading volume consistent with You and Zhang (2009) by scaling the average daily trading volume during the pre- and event periods by the total shares outstanding. I then compare the trading volume by event period across terciles of report length and readability.[24] The results are graphed in Figure 3(a) and 3(b), where Panel A (C) reveals that small investors trade more in both the pre- and event periods when reports are shorter (more readable). In untabulated results, small investor trading volume in the tercile of short reports in Panel A is significantly higher than the trading volume for the tercile of long reports (i.e., *WORDS*) during both the pre- and event periods ($p = .06$ and $p = .02$, respectively). Small investor trading volume is also higher for more readable reports in Panel C, but the differences are insignificant.

In contrast to small investor trading being associated with shorter reports, Panels B and D of Figure 3(b) show that large trade volume is actually greater for firms with longer and less readable reports. Specifically, in Panel B the large investor trade is significantly greater for long reports relative to shorter reports in both the pre- and event periods ($p < .01$; untabulated). Similar

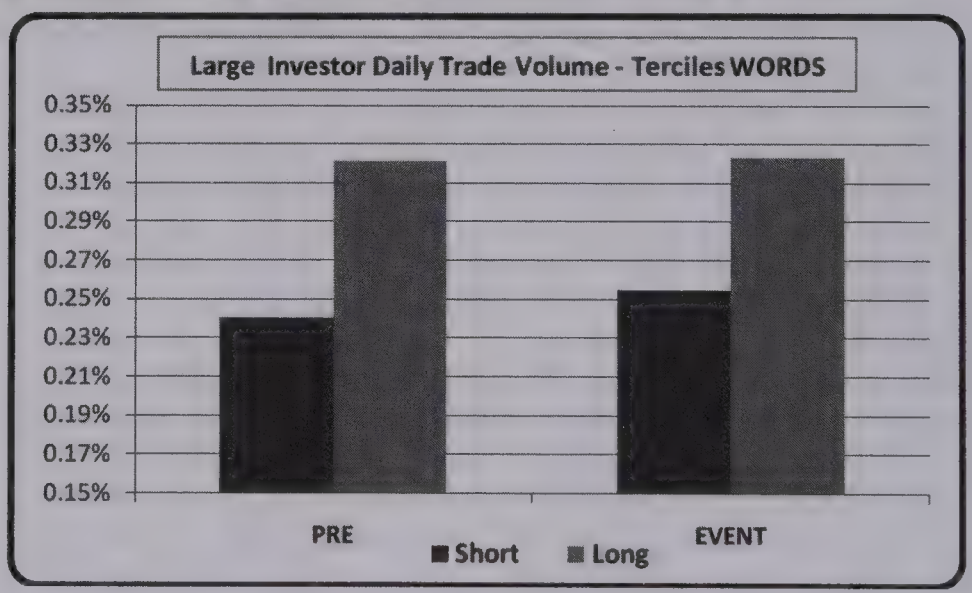
[24] Terciling by year enables me to compare firms within a given year, while parsing out the effects of decimalized trading in the latter periods. I use terciles, based on evidence provided by Lys and Sabino (1992).

FIGURE 3a
Trading by Tercile of Report Complexity

Panel A: Small Investors' Daily Trading Volume by Tercile of *WORDS*



Panel B: Large Investors' Daily Trading Volume by Tercile of *WORDS*



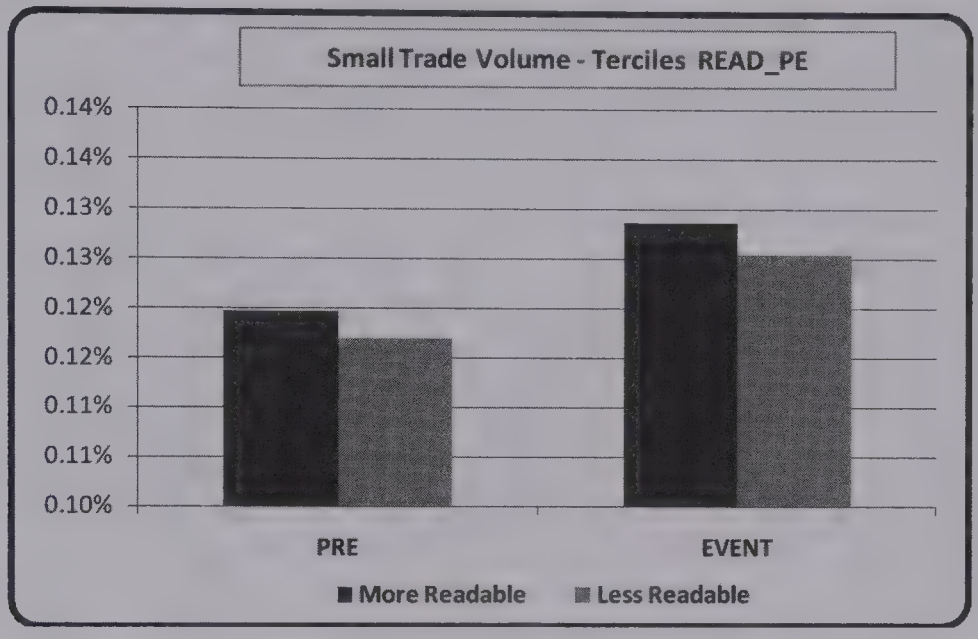
Daily trading volume is the number of shares traded scaled by total number of shares outstanding by investor group (i.e., small or large). Panel A (B) plots the mean of small (large) trade volume based on *WORDS*.

evidence exists in Panel D, wherein large investors trade significantly more during both the pre- and event periods when reports are less readable ($p < .01$ and $p = .03$, respectively).[25] Overall,

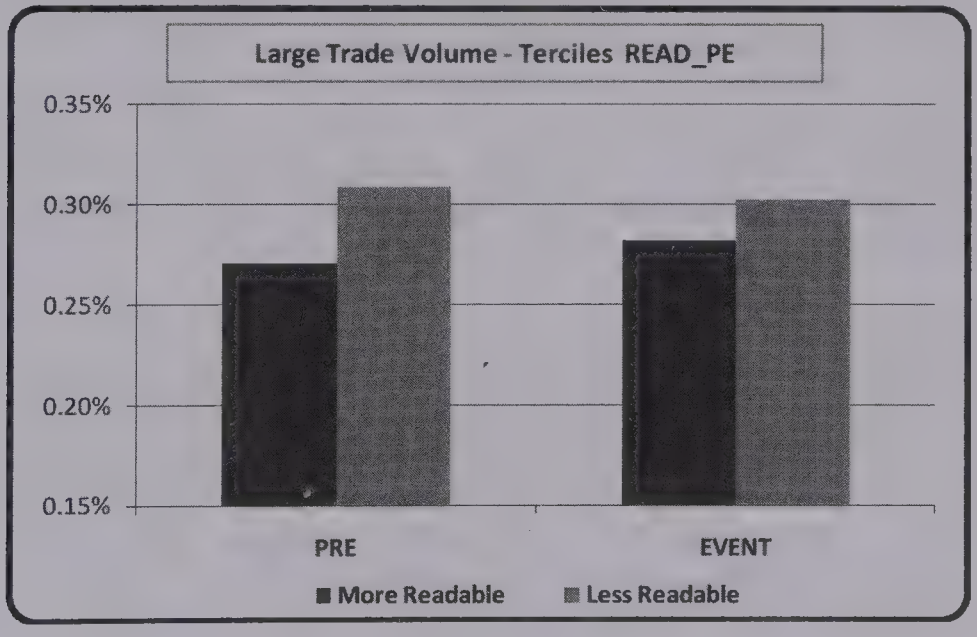
[25] Results using the alternative measures of *WORDS+TABLES* are similar, but more significant for all the relationships

FIGURE 3b
Trading by Tercile of Report Complexity

Panel C: Small Investors' Daily Trading volume by Tercile of *READ_PE*



Panel D: Large Investors' Daily Trading Volume by Tercile of *READ_PE*



Panel C (D) plots the mean of small (large) trade volume based on *READ_PE*. Terciles are created by the appropriate report complexity variable (i.e., *WORDS* or *READ_PE*) within each year to control for the increased proportion of small trade to total trade volume during the latter part of the sample. The middle tercile is eliminated to limit the comparison to more extreme groups.

examined in Figure 2. Similarly, the results for *READ_FOG* are also more significant for large investors, but are insignificant in the predicted direction for small investors.

the evidence in Figure 3(a) and Figure 3(b)and supports the notion that small (large) investors are associated with short (long) reports in both the pre- and event period.[26]

I further examine whether cross-sectional differences in reporting complexity (while holding time constant) affects small and large investor trading behavior in Panel B of Table 5. Consistent with Panel A, the coefficients on *SML* * *COMPLEX* columns (1) through (4) are negative and significant. I also find that the coefficients on *LRG* * *COMPLEX* are negative and significant when the length is measured as the total of both textual and tabulated content in column (2). This finding is consistent with more sophisticated investors being affected by more complex reports issued by certain firms. I continue to find evidence that more complex reports are more likely to lead to lower abnormal trade for small investors relative to large investors, except when examining the effects of *WORDS+TABLES*. However, this lack of difference between large and small investors in column (2) appears more related to large investors being affected by the complexity as measured by the text and tabulated information than by a lack of impact on small investors. Overall, the evidence from Table 5, Panel B is consistent with small investors trading less in the short-window around firms with 10-K filings when firms' reports are more complex.

I also investigate whether small investors are affected by deviations in reporting complexity over time (holding the firm constant) in Table 5, Panel C. Overall, the results are consistent with the earlier panels, although the significance levels of the complexity measures are diminished in all the regressions and the specific effects on *READ_FOG* (*READ_PE*) are insignificant (weakly significant). The weaker findings on *READ_PE* are most likely due to limited variation in the measure over time (see Table 3, Panel C). In summation, the evidence from Panel C supports the notion that increases in report complexity over time have a significant impact on small investors.

Overall, the results in Table 4 and Table 5 provide evidence consistent with a decrease in total trading activity when reports are more complex. This evidence appears due primarily to small investor trading behavior, although there is some evidence that length also has an effect on large investors. Further within-year and within-firm analyses suggest that the association between complex reports and lower abnormal trading appears to be driven by both underlying firm reporting complexity as well as variation in disclosure complexity over time.

Trading Consensus

The negative association between report complexity and small investor trade volume in the previous section is consistent with more complex reports increasing processing costs. I next investigate whether small (large) investors process data in a different way from other small (large) investors when reports are more complex by replacing the dependent variable in Equation (2) with *TRADE_CONS*. Table 6, Panel A provides regression results for small and large investors' consensus trade behavior. Overall, the results are consistent with more complex reports leading to more disagreement among small investors who elect to trade. For example, the coefficients on *SML* * *COMPLEX* in columns (1), (2), (3), and (4) are all negative and significant (all $p < .01$). The coefficients on *LRG* * *COMPLEX* are unexpectedly positive, which is consistent with large investors processing complex disclosures in a similar manner to other large investors when reports are more complex. Finally, I find support in all four columns that more complex reports result in

[26] Although it is possible that report length is a criterion investors use to base their decisions on which firms to follow, investors are more likely to start following a firm based on attributes such as familiarity (Merton 1987).

TABLE 6

The Effects of Reporting Complexity on Small and Large Investor Trade Consensus

Panel A: Pooled Analysis

	Hyp. Sign	*WORDS*		*WORDS+TABLES*		*READ_FOG*		*READ_PE*	
		SMALL	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*
		(1)		(2)		(3)		(4)	
COMPLEX	−	−0.0489***	0.0245	−0.0627***	0.0352	−0.0201***	0.0144	−0.0076***	0.0060
		(−4.33)	(2.50)	(−3.97)	(2.27)	(−5.00)	(3.52)	(−4.04)	(3.27)
AFTEAD	−	0.0008***	0.0001***	0.0001***	0.0001***	0.0008***	0.0001***	0.0002***	0.0007***
		(2.15)	(0.45)	(0.56)	(0.89)	(1.97)	(0.96)	(0.38)	(2.67)
AFTEXD	+	0.0004***	0.0004	0.0002***	0.0009*	0.0005***	0.0004	0.0008***	0.0000
		(0.68)	(1.96)	(1.11)	(4.32)	(0.78)	(1.66)	(1.00)	(0.12)
ABS_RET	+	−0.3333***	−0.4934***	−0.1485***	−0.5127***	−0.3761***	−0.4583***	−0.6080***	−0.2227***
		(−3.64)	(−7.28)	(−3.75)	(−4.65)	(−4.26)	(−7.12)	(−4.97)	(−2.58)
PVOL	?	0.0119***	−0.0124***	0.0040**	−0.0263***	0.0119***	−0.0123***	0.0135***	−0.0138***
		(4.23)	(−4.79)	(2.28)	(−17.04)	(4.04)	(−4.85)	(5.24)	(−4.47)
MV	+/−	−0.0489***	−0.0992***	−0.0173***	−0.1069***	−0.0547***	−0.0955***	−0.0812***	−0.0696***
		(−9.35)	(−25.73)	(−3.46)	(−24.87)	(−11.73)	(−23.39)	(−10.10)	(−8.43)
MTB	?	−0.0137***	−0.0083***	−0.0130***	0.0028	−0.0116***	−0.0093***	−0.0138***	−0.0073*
		(−5.94)	(−2.30)	(−8.20)	(0.97)	(−5.85)	(−2.57)	(−4.75)	(−1.70)
EARNINGS	?	0.0586***	0.0569	0.0031	0.1013*	0.0859***	0.0450*	0.1088***	0.0239
		(1.86)	(2.25)	(0.12)	(1.95)	(2.89)	(1.94)	(3.39)	(0.94)
NY_EARNINGS	?	0.0132***	−0.0536***	0.0027	−0.1121*	0.0136***	−0.0538***	0.0292***	−0.0692***
		(0.70)	(−3.15)	(0.25)	(−3.24)	(0.69)	(−3.05)	(1.28)	(−3.69)
NA_FOLL	−	−0.0251***	−0.0679***	−0.0234***	−0.0576***	−0.0267***	−0.0677***	−0.0173***	−0.0770***
		(−4.13)	(−11.72)	(−6.41)	(−11.56)	(−4.44)	(−11.70)	(−2.33)	(−11.23)
NBSEG	?	−0.0078	0.0116***	0.0060***	0.0113***	−0.0109*	0.0132***	−0.0087	0.0112***
		(−1.23)	(3.67)	(3.09)	(3.15)	(−1.68)	(4.17)	(−1.43)	(3.29)
NGSEG	?	0.0022	−0.0102**	−0.0085*	−0.0071	−0.0001	−0.0080*	−0.0074	−0.0001
		(0.32)	(−2.22)	(−4.60)	(−1.40)	(−0.02)	(−1.76)	(−1.26)	(−0.02)

(continued on next page)

Panel A: Pooled Analysis

	Hyp. Sign	*WORDS*		*WORDS+TABLES*		*READ_FOG*		*READ_PE*	
		SMALL	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*
		(1)		(2)		(3)		(4)	
R^2		55.99%		67.29%		55.50%		51.33%	
Test of (*SML* * *COMPLEX* − *LRG* * *COMPLEX*) < 0		−0.073###		−0.098###		−0.034###		−0.014###	

Panel B: Within-Year Analysis

	Hyp. Sign	*WORDS*		*WORDS+TABLES*		*READ_FOG*		*READ_PE*	
		SMALL	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*
		(1)		(2)		(3)		(4)	
COMPLEX	−	−0.0335***	0.0187	−0.0203***	−0.0001**	−0.0127***	0.0103	−0.0071***	0.0049
		(−16.96)	(8.79)	(−5.18)	(−0.02)	(−17.16)	(11.09)	(−24.68)	(13.39)
CONTROLS			Yes		Yes		Yes		Yes
R^2			59.12%		68.52%		58.90%		55.94%
Test of (*SML* * *COMPLEX*−*LRG* * *COMPLEX*) < 0			−0.052###		−0.020###		−0.023###		−0.012###

(continued on next page)

Panel C: Within-Firm Analysis

	Hyp. Sign	*WORDS*		*WORDS+TABLES*		*READ_FOG*		*READ_PE*	
		SMALL	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*	*SMALL*	*LARGE*
		(1)		(2)		(3)		(4)	
COMPLEX	−	−0.0085**	−0.0002	−0.0127***	0.0036	−0.0027**	0.0000	−0.0001	−0.0006*
		(−1.62)	(−0.08)	(−3.22)	(0.64)	(−1.95)	(0.01)	(−0.21)	(−1.41)
CONTROLS			Yes		Yes		Yes		Yes
R^2			16.46%		11.78%		16.44%		16.41%
Test of (*SML* * *COMPLEX*−*LRG* * *COMPLEX*) < 0			−0.008#		−0.016##		−0.003#		0.001

*, **, *** Indicate two-tailed statistical significance of coefficient estimates at the 10 percent, 5 percent, and 1 percent levels, respectively, when no prediction is given and one-tailed significance when predicted.

#, ##, ### Indicate one-tailed statistical significance of differences between small and large investor coefficients for *COMPLEX* using F-statistics.

This table reports the regression results for small and large investor abnormal trading volume (*TRADE_CONS*).

Panel A reports results from estimating Equation (2) after replacing the dependent variable with *TRADE_CONS*. Panel B reports results from estimating Equation (2) after replacing the dependent variable with *TRADE_CONS* and adding year fixed effects. Panel C reports results from estimating Equation (2) after replacing the dependent variable with *TRADE_CONS* and adding firm fixed effects.

Columns 1, 2, 3, 4 report results when *COMPLEX* is measured as *WORDS*, *WORDS+TABLES*, *READ_FOG*, *READ_PE*, respectively. The reported regressions are stacked regressions of small and large investor groups, where there are two observations (i.e., small and large) for each firm year. Hence, there are a total of 25,542 for each regression, except in column 2 where there are 9,448 observations when *WORDS+TABLES* is the variable of interest. All regressions include a constant term. Regressions are performed with clustered robust standard errors (Rogers 1993) to control for within-firm and -year correlation, except when a specific fixed effect is already included in the model.

greater disagreement among small investors thanamong large investors.[27]

As with abnormal trading volume regressions, the majority of control variables are significant in the expected direction in Table 6, but some of these coefficients also warrant further discussion. Specifically, *PVOL* is significant in different directions for small and large investors, implying that the 10-K resolves prior disagreement for small filers but leads to lower consensus for large investors. However, this interpretation should be taken with caution, as large investors could merely be unraveling their positions based on their different priors about the 10-K filing. The coefficients on both *MV* and *MTB* are significantly negative in all but one of the regressions, suggesting that there is more disagreement among investors of large and growth firms. Finally, the positive coefficients on *EARNINGS* can be interpreted as greater consensus when the underlying profitability is positive, whereas the negative coefficients on *LRG* * *NY_EARNINGS* suggests disagreement among large investors regarding the forward-looking information in the filing.

Consistent with the analysis on trading activity, Table 6 also reports the regression equations including both within-year and within-firm analyses. The coefficients on *SML* * *COMPLEX* are significantly negative in all four of the reported regressions in Panel B. Furthermore, the evidence in all four regressions suggests that small traders are more likely to disagree than their larger, more sophisticated counterparts. Table 6, Panel C provides similar evidence of an association between report complexity and lower consensus trade for small investors. However, the evidence is weaker in all such regressions, and the evidence for *READ_PE*, which has little variation over time, is insignificant. Overall, the evidence in this section reveals that more complex reports appear to be associated with a small but statistically significant decrease in consensus among small investors. For perspective, a 55-page increase results in a decrease in small investor consensus from a median value of 19.9 percent down to 19.5 percent.

In sum, the findings related to consensus are consistent with experimental evidence in Barron et al. (2004), where increases in disclosure lead nonprofessional investors to generate (or infer) private information. However, as opposed to Barron et al.'s (2004) finding that sophisticated professional investors are unaffected by the increased disclosure, there is some evidence that large investors are actually more likely to have increased consensus when more data is provided and the writing is more technical (less readable).

VI. ADDITIONAL ANALYSES

Combined Reporting Attributes

In addition to separately examining the effects of length and readability, I also re-run all analyses including both measures of complexity in the same model. When both length and readability are included in the regressions, the effects of longer reports are similar to those reported in Table 4, Table 5, and Table 6, albeit slightly weaker in a few regressions.[28] However, the effects of readability are insignificant in the predicted direction in all but one regression. This evidence is

[27] Small and large investors are likely to have different levels of prior information. Although these potential differences in private information between large and small investor groups do not affect the within-group comparison, they do make it difficult to interpret the implications of the differential consensus between small and large investors on the abnormal trading results. Specifically, if large investors have more private information, they may be more likely to process information in a similar manner because they have more precise priors leading to greater consensus. As such, it is difficult to determine whether large investors' lower trading reaction to the filings (compared to small investors) is due to common processing or superior processing capabilities.

[28] The coefficients and significance levels on measures of length coefficients are not significantly affected by the inclusion of the readability metrics. Specifically, when the readability measures are included, all the coefficients on *WORDS* and *WORDS+TABLES* in the pooled regressions are significant at the 1 percent level (one-tailed), except when *READ_PE* is examined in the total abnormal volume regressions where the coefficients are significant at the 5 percent and 10 percent levels, respectively. The coefficients on measures of length are quantitatively similar for the within-firm and within-year as well.

consistent with both measures of reporting complexity serving as substitutes, with the effects of longer reports appearing to dominate the readability of the reports when both measures of complexity are analyzed simultaneously.

This analysis is useful in highlighting the differences between this study and a concurrent study by Loughran and McDonald (2010), who also investigate the reporting attributes of the 10-K filings during a similar time frame. While their study focuses primarily on the readability (not length) of financial statements, the two studies overlap in the investigation of small investor trade surrounding the filing. Both studies document that after implementing more sophisticated measures of readability (plain English), improvements in readability lead to increased small investor trading activity. The studies differ in the strength of that relation after simultaneously investigating alternative measures of length.

The discrepancy between the two studies could stem from several differences in composition. For instance, the studies differ in sample size, size cut-offs for small investors, inclusion of certain control variables, trading windows, and the measurement of plain English readability. Perhaps most importantly, differences could stem from research design choices. Specifically, Loughran and McDonald (2010) investigate a change in abnormal trading but include level-control variables. Hence, any deviations in firm characteristics (e.g., size, complexity) or report characteristics (e.g., length) over time are uncontrolled in their analysis. Additionally, as opposed to the abnormal volume measure used in this study, they use a dependent variable that is not de-trended to account for the increase in decimalized trading during the latter part of the sample they investigate. In sum, the evidence across both studies is consistent with small investors being affected by less readable reports, but differences in proprietary measures of plain English readability and research design choices make it difficult to directly compare differences in significance of that association when accounting for report length.

Short-Term versus Long-Term Trading Effects

As previously discussed, the evidence in Figure 1 suggests that the majority of trading occurs in the five-day window surrounding the filing. As such, this study's primary focus is on the effects report complexity on investor trading activity in the short window around the filing. The underlying assumption is that, as processing costs increase, small investors would continue to follow certain stocks but would elect to delay their trades either to give them time to process the reports or to time their trades to coincide with easier to access information events.

To determine whether the trading effects documented above are short-term or long-term in nature, I re-run the analyses in Table 4, Table 5, and Table 6 after increasing the trading window to 10- and 20-day periods. When the event window is lengthened to 10 days, the coefficients on small investors remain negative in the abnormal volume regressions. However, the coefficients and significance is diminished on most tests and become insignificant at conventional levels in about half the regressions. When the windows are expanded to 20 days, the significance on the all the coefficients is further diminished. In contrast to the abnormal volume regressions, the results on *TRADE_CONS* are relatively unaffected by lengthening the windows, as all the coefficients remain negative and significant at the 1 percent level for both the 10- and 20-day windows for all measures of report complexity. Overall, the evidence suggests that while the effects of report complexity are observable in longer windows, they are most pronounced in the short window around the filing.

Other Information Releases

Concurrent Information Releases

As with any event study, it is possible that other information released during the event

window could affect the results. To provide additional assurance that the earnings announcements released in the short window surrounding the 10-K filing are not affecting the results, I eliminate all observations where *AFTEAD* is less than or equal to five days and find that the results are unaffected by removing these observations (untabulated). It is possible that other concurrent information releases may also affect trading behavior. However, in this setting, concurrent information releases would likely act as a filtering mechanism that would bias against finding an effect on information processing.

Prior Information Releases

During the latter part of the sample period investors likely have greater access to alternative sources of information (e.g., financial blogs) that may pre-empt the information in the 10-K filing. These alternative sources of information released prior to the filing could lead to increased trade during the control window and decreased trade during the event window. Due to the trend in report length, this possibility could provide a potential alternative explanation for the findings in this study. To rule out this alternative, I examine variations in the ratio of daily average trading volume during the pre-period divided by the event period by year during the sample period. I find that the ratio of pre-period/event period trading behavior significantly decreases for small investors from the beginning to the end of the sample period ($p < .01$; untabulated), while there is no significant change in the proportion of large investor pre-volume to event period trade. In summary, there is no evidence that increased access to information by either small or large investors during the latter part of the sample period affects trading behavior in a manner that would bias toward the findings.

MD&A Reporting Complexity

In my primary analyses, I examine the length of the entire 10-K document including MD&A, footnotes, contracts, and other exhibits. However, it is possible that the effects are driven by certain parts of the document, such as the MD&A section. To address this issue, I obtain MD&A measures from Li (2009) and examine the MD&A section separately. I find only weak evidence of an effect of MD&A-specific reporting complexity on investor trading, which suggests that the complexity of the MD&A section is not driving the results.

VII. CONCLUSION

The evidence in this study is consistent with more complex filings being too costly for some investors to process. Specifically, I find evidence that more complex filings are associated with reduced trading activity and lower consensus for small investors, but have only a limited impact on large investors. I perform additional within-firm and within-year analyses and find similar results, which suggest that the association between more complex reports and lower abnormal trading appears driven by both relative firm reporting complexity as well as deviations in reporting complexity over time. Additional analysis reveals that the effects of readability and length appear to act as substitutes, with longer reports appearing to dominate the readability of the reports when analyzing both measures of complexity simultaneously.

The evidence provided in this study should prove useful to regulators. For instance, the results highlight the importance of considering processing costs in future disclosure regulation and, therefore, challenge the assumption that more disclosure helps level the playing field between small and large investors. Further, although the SEC has made progress in assisting small investors in gaining electronic access to financial information (Asthana et al. 2004), simply making more data publicly available may not necessarily aid small investors, as the increased cost of processing the additional data makes the 10-K filings too costly for many small investors to process. Finally, initiatives to reduce the processing costs of financial filings may serve as an important factor in ensuring information is impounded quickly into market prices.

APPENDIX A
10-K FILING EDIT PROCEDURES

This appendix provides details of the methods used to download and edit the 10-K filings. I first download all 10-K reports from EDGAR and then edit the filings before calculating the readability scores. I follow the process documented in Li (2008), with some slight modifications.

I delete the heading information that is contained between ⟨SEC-HEADER⟩ and ⟨/SEC-HEADER⟩ and all text that begins with ⟨TABLE⟩ and ends with ⟨/TABLE⟩.[29] I also eliminate all lines that contain ⟨S⟩ or ⟨C⟩ or special characters ⟨...⟩ and ⟨&..⟩, along with all lines that contain strings such as ⟨TEXT⟩, ⟨DOCUMENT⟩, ⟨PAGE⟩, ⟨TYPE⟩, or /PRIVACY-ENHANCED/. Li (2008) eliminates all paragraphs containing these tags. However, these strings often fail to contain end tags making it difficult to delete specific sections. To address this issue, I eliminate only the lines containing this coding.

Further, I convert embedded HTML code to proper ASCII characters and run a Perl package to strip all remaining HTML tags (for further information see http:/www.ascii.cl/htmlcodes.htm). All ".jpg," ".pdf," and ".gif" files are eliminated. Finally, I delete all paragraphs (defined as two lines containing text with a blank line before and after) with more than 50 percent non-alphabetic characters to ensure that all tables, tabulated texts, and financial statements are eliminated. Although this approach is not perfect in eliminating all numeric or HTML code, my review of a large number of these filings indicates that the code does a reasonable job at editing the text. Further, the En Fathom module with Perl attempts to distinguish words from acronyms and other characters, which should minimize any noise attributable to remaining HTML code. There is no reason to believe that any remaining HTML code biases the findings in this study.

Word and Table Counts

I run the En Fathom module in Perl to identify the number of words in the document. To count the number of cells included in the document, I locate all of the HTML tags that correspond to the syntax ⟨TABLE or ⟨table (prior to their deletion as documented above). The number of cells within a table is then computed by counting the number of ⟨TD or ⟨td tags found within a table. Finally, I sum the number of cells across all tables and add that count to the number of words in the document to create the *WORDS+TABLES* variable.

Financial Statements Included by Reference

Finally, to ensure that the results in this study are not affected by firms that include their financial statement by reference, I use a PERL script to identify these firms and find that the results are unaffected by eliminating these observations (untabulated).

APPENDIX B
CALCULATION OF FOG AND PLAIN ENGLISH READABILITY MEASURES

This appendix explains the details of measuring length and creating readability indices after the text has been edited (as described in Appendix A).

---

[29] I follow Li (2008) and delete tables "that begin with ⟨TABLE⟩ and end with ⟨/TABLE⟩." However, I find that certain reports contain significant amounts of text between table tags. For example, in Lowe's Companies 1994 10-K (http://www.sec.gov/Archives/edgar/data/60667/0000060667-94-000015.txt), large portions of text, including the entire MD&A section, are eliminated under this approach. Hence, I repeat the entire analysis without deleting the information between the table tags (but keeping the requirement to delete all paragraphs with more than 50 percent non-alphabetic characters). The results are quantitatively similar using both approaches.

Fog Readability

After the file has gone through the edit process described in Appendix A, the text is then analyzed using the En Fathom package in Perl. This package calculates a variety of textual statistics, including number of characters, number of words, percent of complex words, number of sentences, number of text lines, number of paragraphs, syllables per word, and words per sentence. Based on these statistics, the package produces the Fog readability index.

Plain English Readability

I compute the plain English readability score used in this study by identifying the specific plain English problems highlighted by the SEC and matching them to the specific variables identified by a proprietary text-editing program called StyleWriter (http://www.stylewriter-usa.com).

Plain English Writing Factors

SEC Plain English Problems	StyleWriter Plain English Measure
Passive voice	Passive verbs
Weak/Hidden verbs	Hidden verbs
Superfluous words	Overwriting (overuse of qualifying words)
Legal and financial jargon	Legal words and Jargon/Abstract words
Numerous defined terms	NA
Abstract words	Jargon/Abstract words
Unnecessary details	Tautologies/Overwriting
Long sentences	Number of words/Avg. sentence length
Unreadable design and layout	NA

This software is designed as an add-in editing program to be used in text processing software and, thus, requires the document to be opened in a "word processor" before the document can be analyzed. This limitation results in several additional processing requirements.

First, each document must be opened in text-processing software and then individually analyzed by StyleWriter. I automate this process using a keyboard/mouse click macro to open and score each document, but the need to open and analyze each document individually severely limits the number of documents that can be efficiently processed. Second, the Perl output in .txt format has a "¶" symbol (i.e., carriage return) at the end of each line. This means StyleWriter sees each line as a paragraph, which invalidates several of the plain English results. Hence, prior to processing the documents, I use a Perl script to replace all ¶ with a space and keep all instances where there were two consecutive ¶ symbols (i.e., actual paragraph break). Finally, the program is designed to write out limited summary statistics to allow writers to improve their writing style. However, as originally designed, the program does not output the frequency of all plain English violations. Therefore, I worked with StyleWriter programmers to modify the program to enable me to export the relevant plain English violations for each document.

REFERENCES

Allee, K., N. Bhattacharya, E. Black, and T. Christensen. 2007. Pro forma disclosure and investor sophistication: External validation of experimental evidence using archival data. *Accounting, Organizations and Society* 32 (April): 201–222.

Asthana, S., and S. Balsam. 2001. The effect of EDGAR on the market reaction to 10-K Filings. *Journal of Accounting and Public Policy* 20 (December): 349–372.

———, ———, and S. Sankaraguruswamy. 2004. Differential response of small versus large investors to 10-K

filings on EDGAR. *The Accounting Review* 79 (3): 571–589.

Atiase, R. K. 1985. Predisclosure information, firm capitalization and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–36.

Bamber, L. S. 1987. Unexpected earnings, firm size, and trading volume around quarterly earnings announcements. *The Accounting Review* 62 (July): 510–532.

———, and Y. Cheon. 1995. Differential price and volume reactions to accounting earnings announcements. *The Accounting Review* 70 (July): 417–441.

———, O. Barron, and T. Stober. 1997. Trading volume and different aspects of disagreement coincident with earnings announcements. *The Accounting Review* 72 (October): 575–597.

———, ———, and ———. 1999. Differential interpretations and trading volume. *Journal of Financial and Quantitative Analysis* 34 (September): 369–386.

Barber, B., T. Odean, and N. Zhu. 2009. Do retail traders move markets? *Review of Financial Studies* (January): 151–186.

Barclay, M., and J. Warner. 1993. Stealth trading and volatility: Which trades move prices? *Journal of Financial Economics* (December): 281–305.

Barron, O. 1995. Trading volume and belief revisions that differ among individual analysts. *The Accounting Review* 70 (October): 581–597.

———, D. Byard, and C. Enis. 2004. Leveling the information playing field. *Review of Accounting and Finance* 3 (4): 21–46.

———, D. Harris, and M. Stanford. 2005. Evidence that investors trade on private event period information around earnings announcements. *The Accounting Review* 80 (2): 403–421.

Beaver, W. H. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 6 (Supplement): 67–92.

Bhattacharya, N. 2001. Investors' trade size and trading responses around earnings announcements: An empirical investigation. *The Accounting Review* 76 (2): 221–244.

———, E. Black, T. Christensen, and R. Mergenthaler. 2007. Who trades on pro forma earnings information? *The Accounting Review* 82 (3): 581–619.

Bloomfield, R. 2002. The "Incomplete Revelation Hypothesis" and financial reporting. *Accounting Horizons* 16 (September): 233–243.

———. 2008. Discussion of: Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* (August): 248–252.

———, and P. Fischer. 2009. Disagreement and the cost of capital. Working paper, Cornell University and The Pennsylvania State University.

Botosan, C. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72 (July): 323–349.

Bouwman, M., P. Frishkoff, and P. Frishkoff. 1987. How do financial analysts make decisions? A process model of the investment screening decisions. *Accounting, Organizations and Society* 12 (1): 1–30.

Chambers, A., and S. Penman. 1984. Timeliness of reporting and the stock price reaction to earnings announcements. *Journal of Accounting Research* 22 (Spring): 21–47.

Collins-Thompson, K., and J. Callan. 2005. Predicting reading difficulty with statistical language models. *Journal of the American Society for Information Science and Technology* 56 (November): 1448–1462.

Cready, W., and P. Mynatt. 1991. The information content of annual reports: A price and trading response analysis. *The Accounting Review* 66 (April): 291–312.

———, and D. Hurtt. 2002. Assessing investor response to information events using return and volume metrics. *The Accounting Review* 77 (4): 891–909.

Davis, A., J. Piger, and L. Sedor. 2007. Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases. Working paper, University of Oregon.

Firtel, K. 1999. Plain English: A reappraisal of the intended audience of disclosure under the Securities Act of 1933. *Southern California Law Review* 72 (March): 851–897.

Francis, J., K. Schipper, and L. Vincent. 2002. Expanded disclosures and the increased usefulness of earnings announcements. *The Accounting Review* 77 (3): 515–546.

Griffin, P. A. 2003. Got information? Investor response to form 10-K and form 10-Q EDGAR filings. *Review of Accounting Studies* 8 (December): 433–460.

Grossman, S., and J. Stiglitz. 1980. On the impossibility of informationally efficient markets. *The American*

Economic Review 70 (June): 393–407.

Gunning, R. 1952. *The Technique of Clear Writing*. New York, NY: McGraw-Hill International Book Co.

Hirst, D. E., and P. E. Hopkins. 1998. Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research* 36 (Supplement): 47–75.

Hopkins, P. 1996. The effect of financial statement classification of hybrid financial instruments on financial analysts' stock price judgments. *Journal of Accounting Research* (Supplement): 33–50.

Huber, P. J. 1967. The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, CA.

Hunton, J., and R. McEwen. 1997. An assessment of the relation between analysts' earnings forecast accuracy, motivational incentives, and cognitive information search strategy. *The Accounting Review* 72 (October): 497–516.

Jones, M., and P. Shoemaker. 1994. Accounting narratives: A review of empirical studies of content and readability. *Journal of Accounting Literature* 13: 142–185.

Kandel, E., and N. Pearson. 1995. Differential interpretation of public signals and trade in speculative markets. *The Journal of Political Economy* 103 (August): 831–872.

Karpoff, M. 1986. A theory of trading volume. *The Journal of Finance* 41 (December): 1069–1087.

Kim, O., and R. Verrecchia. 1991. Trading volume and price reactions to public announcements. *Journal of Accounting Research* 29 (Autumn): 302–321.

———, and ———. 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics* 17 (January): 41–67.

———, and ———. 1997. Pre-announcement and event period private information. *Journal of Accounting and Economics* 24 (December): 395–419.

Kross, W., and D. Schroeder. 1984. An empirical investigation of the effect of quarterly earnings announcement timing on stock returns. *Journal of Accounting Research* 22 (Spring): 153–176.

Kyle, A. 1985. Continuous auction and insider trading. *Econometrica* (November): 1315–1335.

Lang, M., and R. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71 (4): 467–492.

Lee, C., and M. J. Ready. 1991. Inferring trade direction from intraday data. *The Journal of Finance* 46 (June): 733–746.

———. 1992. Earnings news and small traders: An intra-day analysis. *Journal of Accounting and Economics* 15 (June/September): 265–302.

———, and B. Radhakrishna. 2000. Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets* 3 (2): 83–111.

Lehavy, R., F. Li, and K. Merkley. 2009. The effect of annual report readability on analyst following and the properties of their earnings forecasts. Working paper, University of Michigan.

Levitt, A. 1997. Corporate finance in the information age. Remarks delivered to the Securities Regulation Institute, San Diego, California, January 23. Available at: http://www.sec.gov/news/speech/speecharchive/1997/spch135.txt.

Li, E., and K. Ramesh. 2009. Market reaction surrounding the filing of periodic SEC reports. *The Accounting Review* 84 (4): 1171–1208.

Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* (August): 221–247.

———. 2009. The information content of the forward-looking statements in corporate filing—A naive Bayesian machine learning approach. Working paper, University of Michigan.

Loughran, T., and B. McDonald. 2010. Measuring readability in financial text. Working paper, University of Notre Dame.

Lys, T., and J. Sabino. 1992. Research design issues in grouping-based tests. *Journal of Financial Economics* 32 (December): 355–387.

Maddala, G. 2001. *Introduction to Econometrics*. 3rd edition. New York, NY: John Wiley & Sons.

Maines, L., and L. McDaniel. 2000. Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. *The Accounting Review* 75 (2): 179–207.

Mikhail, M., B. Walther, and R. Willis. 2007. When security analysts talk, who listens? *The Accounting Review* 82: 1227–1253.

Merton, R. 1987. A simple model of capital market equilibrium with incomplete information. *The Journal of Finance* 42: 483–510.

Meulbroek, L. 1992. An empirical analysis of illegal insider trading. *The Journal of Finance* (December): 1661–1699.

Morse, D. 1981. Price and trading volume reaction surrounding earnings announcements: A closer examination. *Journal of Accounting Research* (Autumn): 374–383.

Muresan, G., M. Cole, C. Smith, L. Liu, and N. Belkin. 2006. Does familiarity breed content? Taking account of familiarity with a topic in personalizing information retrieval. *Proceedings of the Hawaii International Conference on System Sciences* (HICSS-39).

Nelson, K., and A. Pritchard. 2007. Litigation risk and voluntary disclosure: The use of meaningful cautionary language. Working paper, Rice University.

Paredes, T. 2003. Blinded by the light: Information overload and its consequences for securities regulation. *Washington University Law Quarterly* 81 (Summer): 417–485.

Plumlee, M. 2003. The effect of information complexity on analysts' use of that information. *The Accounting Review* 78 (1): 275–296.

Radin, A. 2007. Have we created financial statement disclosure overload? Available at: http://www.nysscpa.org/cpajournal/2007/1107/perspectives/p6.htm.

Rogers, W. 1993. *Regression Standard Errors in Clustered Samples*. Stata Technical Bulletin 13. College Station, TX: Stata Press.

Securities and Exchange Commission. 1998. A plain English handbook: How to create clear SEC disclosure. SEC Office of Investor Education and Assistance. Available at: http://www.sec.gov/pdf/handbook.pdf.

———. 2000a. Fair Disclosure Regulation FD. Available at: http://www.sec.gov/rules/final/33-7881.htm.

———. 2000b. Final rule: Disclosure of order execution and routing practices. Available at: http://www.sec.gov/rules/final/34-43590.htm.

———. 2010. The investor's advocate: How the SEC protects investors, maintains market integrity, and facilitates capital formation. Available at: http://www.sec.gov/about/whatwedo.shtml.

Shanthikumar, D. 2004. Small trade reactions to consecutive earnings surprises. Dissertation, Stanford University.

Tetlock, P. 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance* 62 (3): 1139–1168.

———, M. Saar-Tsechansky, and S. Macskassy. 2008. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance* 63 (3): 1437–1467.

You, H., and X. Zhang. 2009. Financial reporting complexity and investor under-reaction to 10-K information. *Review of Accounting Studies* 14 (4): 559–586.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48: 817–883.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 2145–2167

American Accounting Association
DOI: 10.2308/accr.2010.85.6.2145

Can Identifying and Investigating Fraud Risks Increase Auditors' Liability?

Andrew B. Reffett
Miami University

ABSTRACT: Legal scholars and accounting practitioners have expressed concern that the U.S. legal system might, in cases of undetected fraud, effectively penalize auditors for identifying and investigating fraud risks (AICPA 2004; Coffee 2004; Golden et al. 2006). This study draws on counterfactual reasoning theory to provide experimental evidence indicating that this concern is warranted. Consistent with counterfactual reasoning theory, I find that evaluators in a natural (i.e., between-participants) environment are *more* likely to hold auditors liable for failing to detect fraud when the auditors investigated for the perpetrated fraud, relative to when the auditors did not investigate for the fraud. Findings from a less natural within-participants experiment, however, indicate that the between-participant findings likely were unintentional: That is, the same evaluators, when later asked to judge alternative levels of fraud investigation simultaneously, are *less* likely to hold auditors liable for failing to detect fraud when the auditors investigated for the fraud, relative to when the auditors did not investigate for the fraud.

Keywords: *fraud risks; auditor liability; counterfactual reasoning theory; negative affect.*

Data Availability: *Data and experimental materials used in this study are available upon request.*

I. INTRODUCTION

Auditing standards require auditors to identify fraud risks during the planning stage of their audits and then design audit procedures to investigate the identified risks (AICPA 2002, 2006). While identifying and investigating fraud risks should help auditors detect fraud, several publications suggest such actions could increase auditors' liability in the event of an audit failure. In particular, a SAS No. 99 practice aide advises auditors against documenting the specific

This paper is based on my dissertation completed at the University of Illinois. I gratefully acknowledge the helpful insights of my committee members, Clifton Brown (chair), Joel Pike (director of research), Mark Peecher, and Jennifer Robbennolt. This paper also has benefited from valuable comments from Brian Ballou, Rachel Birkey, Billy Brewster, Tim Eaton, Brooke Elliot, Anne Farrell, Jonathan Grenier, Dan Heitger, Michael Hughes, Steven Kachelmeier, Kathryn Kadous, Lisa Koonce, Jennifer Mueller, David Piercey, James Preston, Steve Smith, two anonymous reviewers, and the accounting workshop participants at the University of Maine, Miami University, Indiana University, Oregon State University, Notre Dame University, University of Evansville, University of New Hampshire, University of Texas, and University of Illinois. Finally, I acknowledge the financial support of the Richard D. and Anne Marie Irwin fellowship.

Editor's note: Accepted by Kathryn Kadous.

Submitted: August 2008
Accepted: April 2010
Published Online: November 2010

fraud risks they identify during the required fraud brainstorming sessions because doing so may provide unnecessary litigation exposure to the firm (AICPA 2004). In addition, a recent forensic accounting guide, created by a partner in a Big 4 forensic accounting practice, warns that enhancing audit procedures to meet the public's heightened expectations of auditors to detect fraud may further raise expectations and the accompanying risk of litigation (Golden et al. 2006). Similar concern has also been expressed in the law literature. Specifically, Coffee (2004) claims that auditors' best defense in cases of undetected fraud is ignorance, such that auditors have an incentive not to inquire too closely—lest one acquire information to put one on notice.

The objective of this study is to leverage relevant psychology theory to empirically test whether the above concerns are warranted. Counterfactual reasoning theory indicates that investigating fraud risks likely can increase auditors' liability in the event of an audit failure due to undetected fraud. In particular, based on prior research examining the antecedents and consequences of counterfactual reasoning (Kahneman and Miller 1986; Miller and McFarland 1986; Macrae 1992; Sherman and McConnell 1995), I predict that evaluators in cases of undetected fraud will experience more intense counterfactual thoughts of what the auditors could have done differently to detect the fraud and will judge auditors more harshly when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud.

To test these predictions, I conduct an experiment in which participants serving in the role of individual jurors evaluate auditors who failed to detect fraud during a recent audit. I manipulate whether the auditors performed audit procedures to investigate a specific fraud risk related to the perpetrated fraud. After reading the case, participants decide, first in a between-participants environment and then in a within-participants environment, whether the auditors were negligent and therefore liable for plaintiff losses.

Several factors suggest that jurors in real-world auditor negligence lawsuits make auditor liability decisions in environments more similar to a between-participants environment than a within-participants experimental environment. Specifically, the unique risk factors surrounding individual real-world audits and jurors' general lack of auditing experience and limited knowledge of audit procedures suggest that jurors in real-world auditor negligence lawsuits are not able to compare alternative levels of auditor investigation across identical circumstances. Thus, participants' between-participants responses provide a more realistic indication of jurors' assessments of auditor liability in cases of undetected fraud than participants' within-participants responses. This view is consistent with prior experimental studies of jurors' assessments of auditor liability, which test their hypotheses in between-participants environments (Kadous 2000, 2001; Lowe et al. 2002; Clarkson et al. 2002).

Within-participants responses nevertheless are informative. Specifically, the salience of manipulated factors in within-participants designs enables participants to identify and correct inconsistencies between their natural (i.e., between-participants) judgments and their beliefs about how their judgments should vary with levels of a manipulated factor. In this sense, participants' responses in within-participants environments provide an indication as to whether their between-participants judgments are intentional (Kahneman and Tversky 1996).

Consistent with counterfactual reasoning theory, I find that evaluators in a between-participants environment experience more intense thoughts of what the auditors could have done differently to detect the fraud and are *more* likely to hold auditors liable for plaintiff losses when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. The observed between-participants decision pattern, however, appears to be unintentional, as I find that the same evaluators, when subsequently placed in a within-participants environment, are *less* likely to hold auditors liable for plaintiff losses when the

auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud.

This study makes several important contributions to the accounting literature. Scholars and practitioners have expressed concern that the U.S. legal system might, in cases of undetected fraud, effectively penalize auditors for investigating fraud risks (Coffee 2004; Golden et al. 2006). This study informs their arguments with relevant psychology theory and provides evidence to indicate whether such concern is warranted. Results indicate that performing audit procedures to investigate specific fraud risks can increase jurors' assessments of auditor liability in cases of undetected fraud. This study therefore supports the validity of recently expressed concerns regarding the potential legal hazards of investigating fraud risks.

In addition, by examining evaluator judgment in both between- and within-participants environments, this study advances the current understanding of evaluators' assessments of auditor liability in cases of undetected fraud. The study's within-participants results demonstrate that evaluators, when able to see and compare alternative levels of fraud investigation simultaneously, understand that higher levels of fraud investigation indicate higher levels of auditor care and, therefore, imply lower levels of auditor liability. The between-participants results, in contrast, demonstrate that evaluators have difficulty applying this standard in the more natural setting where alternative levels of auditor investigation under identical circumstances likely do not exist and are, therefore, unavailable for comparison.

Improving evaluators' liability decisions requires that evaluators understand normative standards and are willing to apply those standards (Wegener et al. 2000). The within-participants results demonstrate that evaluators in cases of undetected fraud meet these two criteria. This study therefore suggests that the observed between-participants decision pattern potentially could be reversed. Additional empirical research could be directed at identifying an effective debiasing method to reverse evaluators' unintentional tendency, when evaluating individual cases of undetected fraud, to judge auditors who performed audit procedures to investigate for the fraud more harshly than auditors who did not investigate for the fraud.

Section II develops theory and presents the study's hypotheses. Section III outlines the experimental design and method. Section IV presents results. The final section summarizes, identifies directions for future research, and discusses limitations.

II. THEORY AND HYPOTHESES

Background

Taking actions that increase the quality of an audit should, based on liability laws, reduce auditor liability in the event of an audit failure (Causey and Causey 1991; Kadous 2000). Thus, economic models of auditor liability assume that taking actions to increase the quality of an audit will reduce auditor liability in the event of an audit failure (Smith and Tidrick 1998; Radhakrishnan 1999). Economic models of auditor liability therefore imply that performing audit procedures to investigate fraud risks (assuming such actions increase audit quality) will reduce auditor liability in the event of an audit failure.

Empirical research, however, suggests that evaluators often lack the knowledge and/or ability to make decisions that are consistent with legal standards (Robbennolt et al. 2003). Information disclosed during a trial therefore does not always affect juror decision-making in a manner that is consistent with legal standards. In an audit litigation setting, Kadous (2000) finds that juror standards of auditor care increase with the severity of the consequences associated with an audit failure (despite legal standards that indicate juror standards of care should remain independent of consequence severity). This finding suggests that jurors and other lay evaluators lack a clear understanding of what auditors should be held responsible for and instead rely on the single case they are presented to establish the appropriate standard of care. As a result, taking actions that

increase the quality of an audit might not decrease auditors' liability in the event of an audit failure (Kadous 2000). With respect to fraud, performing audit procedures to investigate fraud risks might not decrease and could, as Coffee (2004) and Golden et al. (2006) suggest, increase auditors' liability in the event of an audit failure due to undetected fraud.

Below, I review prior research examining the antecedents and consequences of counterfactual reasoning and then apply the main findings to the current study's audit litigation setting. As I discuss below, prior psychology research suggests that counterfactual reasoning will significantly affect evaluators' assessments of auditor liability in cases of undetected fraud. Counterfactual reasoning theory, therefore, provides an appropriate basis for developing expectations regarding the effects of investigating fraud risks on evaluators' assessments of auditor liability in cases of undetected fraud.

Antecedents of Counterfactual Reasoning

In the context of harmful events, counterfactual reasoning refers to imagining how prior outcomes could have been avoided (Roese 1997; Alicke et al. 2008). Prior research on accidental injury/death litigation indicates that juror assessments of defendant liability increase with jurors' counterfactual thoughts (Miller and McFarland 1986; Bothwell and Duhon 1994). Analogously, in cases of undetected financial statement fraud, evaluators' assessments of auditor liability likely will increase with the extent of their counterfactual thoughts as to what the auditors could have done differently to detect the fraud.

The extent to which evaluators engage in counterfactual reasoning following negative outcomes depends on the perceived mutability of the prior events (Kahneman and Miller 1986). Specifically, the easier it is for evaluators to imagine more favorable alternative outcomes to prior events, the more evaluators will engage in counterfactual reasoning. As I discuss below, various factors affect the perceived mutability of prior events (Roese 1997). Consequently, in cases of undetected fraud, the extent to which evaluators engage in counterfactual reasoning (and therefore the effect of counterfactual reasoning on evaluators' assessments of auditor liability) likely depends on the presence or absence of several factors that affect the perceived mutability of prior events.

One factor that affects the perceived mutability of prior events is outcome closeness. Specifically, it is easier to mentally alter prior negative events and imagine more favorable alternative outcomes when prior events occur by narrow as opposed to wide margins (Kahneman and Miller 1986). Close outcomes (i.e., near misses) facilitate the imagination of alternative outcomes because it is easier to imagine alternative outcomes when one can identify specific occurrences that, if altered, would have changed the outcome of the event in question. Since identifying such critical occurrences is easier when negative outcomes occur by narrow, as opposed to wide, margins, it is easier to imagine alternative outcomes following near misses (Kahneman and Miller 1986).

In addition to close outcomes, prior research also indicates that actions are easier to mutate than non-actions. Consequently, counterfactual reasoning increases following actions as opposed to non-actions (Gleicher et al. 1990; Sherman and McConnell 1995). Prior studies offer several explanations for the effect of action (as opposed to inaction) on the extent of evaluators' counterfactual thoughts (Landman 1987). One explanation relevant to this study is that action directs evaluators' attention to, and therefore highlights individuals' causal roles in, prior negative outcomes (Creyer and Gurhan 1997). This increases the belief that altering individuals' behavior likely would result in a more favorable outcome, which makes it easier to imagine such an alternative outcome. Thus, evaluators' counterfactual thoughts of what individuals could have done differently to avert a negative outcome tend to increase following actions as opposed to non-actions (Creyer and Gurhan 1997).

Both factors discussed above suggest that evaluators in cases of undetected fraud will experience greater counterfactual thoughts of what the auditors could have done differently to detect the fraud when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. Specifically, I predict that, in cases of undetected fraud, evaluators will perceive auditors who performed audit procedures to investigate for the fraud to have been closer to detecting the fraud than auditors who did not investigate for the fraud. Thus, based on prior research examining the effect of outcome closeness on counterfactual reasoning, I expect evaluators' thoughts of what the auditors could have done differently to detect the fraud to increase when the auditors performed audit procedures to investigate for the fraud. In addition, prior research examining the effect of actions (versus non-actions) on counterfactual reasoning also indicates that evaluators' thoughts of what the auditors could have done differently to detect the fraud will increase when the auditors performed audit procedures to investigate for the fraud (i.e., took action), relative to when the auditors did not investigate for the fraud (i.e., did not take action).

When examining the extent of evaluators' counterfactual thoughts, prior research considers several factors, including both frequency and intensity (Roese 1994; Sanna and Turley-Ames 2000). Sanna and Turley-Ames (2000) find that counterfactual intensity explains variance in affective reactions beyond that explained by counterfactual frequency, while counterfactual frequency does not account for variation beyond that of counterfactual intensity. Accordingly, I primarily consider counterfactual intensity as opposed to counterfactual frequency. In particular, I expect evaluators in cases of undetected fraud to experience *more intense* counterfactual thoughts of what the auditors could have done differently to detect the fraud when the auditors performed audit procedures to investigate for the fraud.

H1: Evaluators in cases of undetected fraud will experience more intense thoughts of what the auditors could have done differently to detect the fraud when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud.

Consequences of Counterfactual Reasoning

As noted above, evaluators' assessments of individuals' responsibility for prior negative outcomes increase with the extent of their counterfactual thoughts (Miller and McFarland 1986; Macrae 1992; Bothwell and Duhon 1994). Prior research demonstrates that the consequences of counterfactual reasoning are widespread, including emotions and causal inferences (Roese 1997). Thus, several affective and cognitive mechanisms could mediate the relationship between counterfactual reasoning and evaluator assessments of individuals' liability for prior negative outcomes. In the following paragraphs, I examine several specific consequences of counterfactual reasoning that I expect to increase evaluators' assessments of auditor liability in cases of audit failure due to undetected fraud.

One frequently cited reason why counterfactual reasoning increases evaluators' liability assessments is that counterfactual reasoning amplifies evaluators' emotional (affective) reactions to prior negative outcomes (Robbennolt and Sobus 1997). Counterfactual reasoning (i.e., thinking about what could have been done differently to avoid a negative outcome) increases the belief that a prior outcome could have been avoided (McMullen et al. 1995; Roese 1997). The belief that a prior negative outcome could have been avoided makes the actual outcome appear more negative and poignant (Roese 1997), which in turn amplifies evaluators' affective responses to the event (Markman et al. 1993; Medvec et al. 1995; Roese and Olson 1995; Alicke et al. 2008). Consequently, evaluators' affective responses to prior negative outcomes increase with the intensity of their counterfactual thoughts (Sanna and Turley-Ames 2000).

Increased affective reactions among evaluators in litigation settings have been found to result in antipathy for defendants and sympathy for plaintiffs (Macrae 1992). Therefore, I expect more intense counterfactual thoughts among evaluators (and thus greater affective reactions) to result in greater antipathy for defendants and sympathy for plaintiffs. Prior research indicates that negative affect directed toward defendants increases the harshness of evaluators' assessments of defendant liability (Macrae 1992). In an audit litigation setting, Kadous (2001) finds that evaluators in cases of undetected fraud use negative affect as a cue to auditor blameworthiness.[1] In addition, Miller and McFarland (1986) report that increased sympathy for plaintiffs results in harsher judgments of defendants. Consequently, in cases of undetected fraud, the combination of increased antipathy for the auditors and increased sympathy for the plaintiff likely will result in more severe judgments of auditor liability. Therefore, I expect more intense counterfactual thoughts (and thus greater auditor antipathy/plaintiff sympathy among evaluators) to result in harsher evaluations of the auditors in cases of undetected fraud.

Prior psychology research also indicates that the effects of counterfactual reasoning extend beyond evaluators' affective reactions (Roese 1997). Therefore, in cases of undetected fraud, the intensity of evaluators' counterfactual thoughts likely affects cognitive assessments that are relevant to the evaluators' ultimate assessments of auditor liability. One such assessment likely influenced by counterfactual reasoning is evaluators' assessment of auditor competence. In particular, when auditors fail to detect fraud, thinking about what the auditors could have done differently to detect the fraud could highlight the apparent necessity of the omitted procedures and, thus, decrease evaluators' assessments of auditor competence. Consequently, in cases of undetected fraud, I expect evaluators' assessments of auditor competence to directly decrease with the intensity of their thoughts of what the auditors could have done differently to detect the fraud.

In addition, prior research suggests that counterfactual reasoning could indirectly decrease evaluators' assessments of auditor competence through evaluators' affective reactions. There are several reasons why evaluators' affective reactions could affect their assessments of auditor competence. First, prior research indicates that information that is linked with an individual's current affective state is more likely to be activated, recalled, and used in cognitive tasks, which leads to congruency between the individual's affective state and her judgments (Bower 1981; Bower and Forgas 2001). Thus, auditor antipathy/plaintiff sympathy could result in a disproportionate weighting of case information that is unfavorable to the auditor and ultimately decrease evaluators' assessments of auditor competence. Second, evaluators could view their affective reactions as information regarding their assessment of auditor competence (Schwarz and Clore 1983; Kadous 2001). Specifically, evaluators could view negative feelings toward the auditors as an indication that the auditors did not perform the audit competently. Based on these reasons, I expect more intense counterfactual thoughts to indirectly decrease evaluators' assessments of auditor competence through affective reactions.

Finally, counterfactual thoughts imply causal connections that can directly affect assessments of blame (Roese 1997). For example, in cases of undetected fraud, thinking about what an auditor could have done differently to detect the fraud indicates that the audit failure would not have occurred if the auditor would have acted differently. In turn, the belief that the audit failure would not have occurred if the auditor would have acted differently implies that the audit failure was caused by the auditors. Such causal implications naturally increase evaluators' assessments of blame (Roese 1997). Thus, beyond the effects of counterfactual reasoning on evaluators' affective

[1] Cornell et al. (2009) report that, by issuing an apology in cases of audit failure, auditors can effectively reduce evaluators' affective reactions to an audit failure, which decreases evaluators' need to assign blame and ultimately reduces the severity of their assessments of auditor negligence. This finding provides further empirical support for the link between evaluators' affective responses and their assessments of auditor negligence in cases of audit failure.

reactions and their assessments of auditor competence, I expect evaluators' assessments of auditor liability to directly increase with the intensity of their counterfactual thoughts.

Summary of the Antecedents and Consequences of Counterfactual Reasoning

Prior research, for several reasons, suggests that evaluators in cases of undetected fraud will experience more intense counterfactual thoughts of what the auditors could have done differently to detect the fraud when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud (H1). As a result, I expect evaluators to be more likely to experience antipathy for the auditors, to be more likely to experience sympathy for the plaintiff, and to have lower assessments of auditor competence when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. Due to these effects, and also due to direct effects of counterfactual reasoning on evaluators' liability assessments, I expect evaluators' between-participants assessments of auditor liability to be more severe when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. Figure 1 presents a model of the hypothesized effects of investigating fraud risks on evaluators' assessments of auditor liability in cases of undetected fraud.

H2: Evaluators' between-participants assessments of auditor liability in cases of undetected fraud will be more severe when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not perform audit procedures to investigate for the fraud.

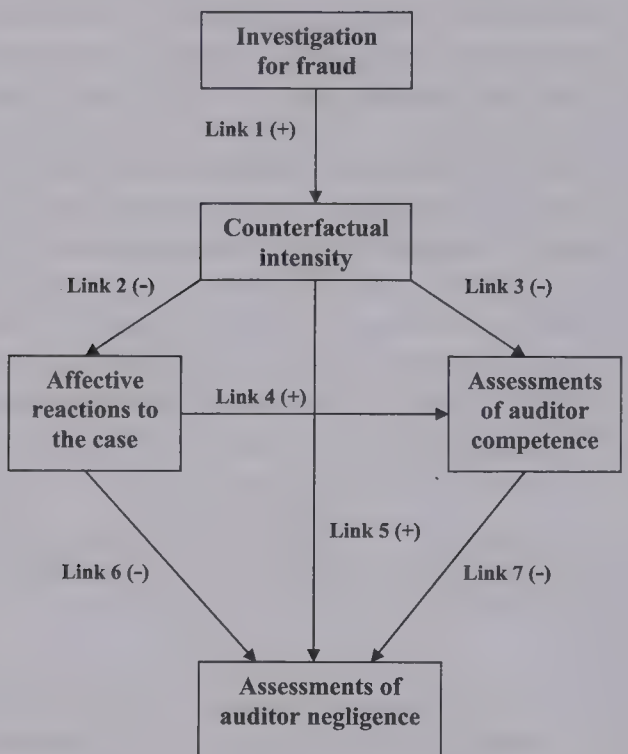
Importantly, while prior research is consistent with H2, the prediction does not lack tension. In studying the effects of counterfactual reasoning on juror assessments of defendant liability, Miller and McFarland (1986) manipulate outcome closeness (which, in turn, positively affects the extent of evaluators' counterfactual thoughts) through a factor that is unrelated to the defendant's level of care. If one holds defendant care constant, theory clearly predicts that increased counterfactual reasoning will result in harsher liability assessments of defendants. In the current study's audit litigation setting, however, the same factor expected to increase the intensity of evaluator counterfactual thoughts (i.e., whether the auditors performed audit procedures to investigate for the fraud) also increases the level of care provided by the auditors. That is, in this setting, auditors provide a higher level of care when the evaluators' counterfactual reasoning is expected to be relatively high, and a lower level of care when their counterfactual reasoning is expected to be relatively low. As such, the higher level of care provided by auditors could offset the effects of increased counterfactual reasoning on evaluators' assessments of auditor liability. Consequently, it is not self-evident that the results of the experiment will support H2.

Within-Participants Assessments of Auditor Liability

In addition to examining evaluators' between-participants assessments of auditor liability, I also examine evaluators' within-participants assessments of auditor liability. The purpose of examining participants' within-participants assessments is to examine whether evaluators' assessments of auditor liability in the more natural between-participants environment are intentional (Kahneman and Tversky 1996). In the following paragraphs, I consider whether evaluators' assessments of auditor liability in a within-participants environment likely vary from their assessments of auditor liability in a between-participants environment.

Hypothesis 2 predicts that evaluators' between-participants assessments of auditor liability in cases of undetected fraud will be more severe when the auditors performed audit procedures to investigate for the fraud. One reason why this prediction might not hold, however, is that evaluators *could* have more favorable assessments of the level of care provided by the auditors when the

FIGURE 1
Hypothesized Model of the Effects of Investigating Fraud Risks on Evaluator Assessments of Auditor Liability



This figure illustrates my proposed model of how performing audit procedures to investigate fraud risks can affect evaluators' assessments of auditor liability in cases of undetected fraud. The parenthetical comment next to each link represents the expected coefficient sign.
"Investigation for fraud" refers to whether the auditors performed audit procedures to investigate for the fraud.
"Counterfactual intensity" refers to the intensity of participants' thoughts of what the auditors could have done differently to detect the fraud.
"Affective reactions to the case" refers to participants' affective reactions to the case. Positive (negative) numbers are indicative of affective reactions that are favorable to the auditor (plaintiff).
"Assessments of auditor competence" refers to participants' assessments of how competent the auditors were in their performance of the audit.
"Assessments of auditor negligence" refers to participants' probability assessments of how likely it was that the auditors were negligent in their performance of the audit.

auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. Nevertheless, I expect more intense counterfactual thoughts of what the auditors could have done differently to detect the fraud, in a between-participants environment, to adversely affect evaluators' assessments of the level of care provided by the auditors.

In a within-participants environment, however, I do not expect more intense counterfactual thoughts to adversely affect evaluators' assessments of the level of care provided by the auditors. The results of a survey by Frank et al. (2001) suggest that non-accountant evaluators believe that auditors who investigate for fraud provide a higher level of care than auditors who did not investigate for fraud, *ceteris paribus*. Therefore, when evaluators in cases of undetected fraud are

placed in an environment where they can directly compare auditors who investigated for the fraud versus auditors who, under *identical circumstances,* did not investigate for the fraud, I expect evaluators to conclude that the auditors who investigated for the fraud provided a higher level of care than the auditors who did not. Thus, I expect evaluators' within-participants assessments of auditor liability to be less severe when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud.

H3: Evaluators' within-participants assessments of auditor liability in cases of undetected fraud will be less severe when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not perform audit procedures to investigate for the fraud.

III. EXPERIMENTAL DESIGN AND METHOD

To test H1 and H2, I employ a between-participants design that manipulates whether, and to what extent, the auditors performed audit procedures to investigate for the fraud (no investigation, low investigation, and high investigation). To test H3, I employ a within-participants design that includes the same manipulations as in the between-participants setting. Thus, to test H1, H2, and H3, I compare participant responses across the following conditions: (1) no investigation; (2) low investigation; and (3) high investigation.

Participants

Two hundred twenty-nine undergraduate students participated in the study. Participants were recruited at two U.S. state universities from a wide array of non-business school courses. To motivate effort, participants were truthfully told that participants who correctly answered at least 85 percent of a series of review questions would be entered in a cash lottery. Prior research finds that verdicts from student participants do not differ from more diverse samples of participants (Bornstein 1999; Zickafoose and Bornstein 1999). Similarly, in an audit litigation setting, Kadous (2001) finds no significant differences in the judgments made by student participants versus participants selected for jury duty. Thus, students appear to be appropriate participants for this study.

Experimental Procedures

Participants were first randomly assigned to experimental conditions (see "Independent Variables" below) and were given an experimental case to read and complete.[2] The case packet provides background information regarding the nature and purpose of financial statements and the external audit process, the details of an audit of a fictional mining company, the details of a financial statement fraud the auditors failed to detect, and the transcript from a subsequent auditor negligence lawsuit. The undetected financial statement fraud pertains to the mining company's intentional understatement of its environmental restoration liability. The mining company perpetrated the fraud by misrepresenting to the auditors the costs they expected to incur to reclaim one of their inactive quarries. The transcript from the lawsuit includes opening and closing statements by both plaintiff and defense attorneys and testimony by equally credentialed expert witnesses for both the plaintiff and defense.

After reading the case, participants make several assessments of auditor liability for plaintiff damages (see dependent variables). Then, participants complete a post-experimental survey (see mediating variables). Finally, participants complete a within-participants questionnaire, in which participants read different scenarios describing how the auditors, in the same circumstance as the

[2] The experimental case is adapted from the cases used in Peecher and Piercey (2008) and Kadous (2000).

auditor described in the first case, performed the audit differently but still failed to detect the fraud.[3] The audit scenarios presented in the within-participants setting correspond to the conditions the participants did not see in the first experimental case. After each scenario, participants make assessments of auditor liability for the auditors in the given scenario.

Independent Variables

I first manipulate whether the auditors performed audit procedures to investigate for the perpetrated fraud. In the no-investigation condition, the auditors do not identify the fraud as a fraud risk and, thus, do not perform audit procedures to investigate for the fraud (no-investigation condition). In the investigation conditions, the auditors perform audit procedures to investigate for the fraud. In one investigation condition, the auditors investigate for the fraud by making inquiries of the company's president (low-investigation condition). In the other investigation condition, the auditors make inquiries of the company president and also hire a geology firm to test the soil from a sample of the company's quarries for hazardous materials that could increase their environmental restoration liability (high-investigation condition). I include two investigation conditions to examine whether the effect of performing audit procedures to investigate fraud risks on evaluators' assessments of auditor liability in cases of undetected fraud varies across different levels of auditor investigation.

Dependent Measures

The experiment examines the intensity of evaluators' counterfactual thoughts of what the auditors could have done differently to detect the fraud. To measure the intensity of participants' counterfactual thoughts, participants are asked on a post-experimental survey how seriously, on a scale from 0 to 100, they thought about what the auditors could have done differently to detect the fraud. Sanna and Turley-Ames (2000) use a similar measure of counterfactual intensity. The intensity of participants' counterfactual thoughts is the dependent variable of interest for H1, and is expected to mediate the effect predicted in H2.

The experiment also examines three measures of evaluators' assessments of auditor liability for plaintiff losses in cases of undetected fraud. The first measure asks participants to indicate whether the auditors were negligent in performing the audit (yes, no). The second measure asks participants to indicate how probable (0–100 scale) it is that the auditors were negligent in performing the audit. Participants who indicate that the auditors were negligent are then asked to indicate the amount of compensatory damages (0–$9 million) they would require the auditors to pay to the plaintiff.

Participants' yes/no decisions of auditor negligence and their probability assessments of auditor negligence are the primary dependent variables of interest for H2 and H3, as those two variables are more direct measures of evaluators' assessments of whether the auditors are liable for plaintiff losses, which is the study's construct of interest. Participants' damage assessments, while a less direct measure of evaluations of negligence, are measured and reported to indicate the potential economic significance of the effects of differences in participants' assessments of auditor negligence across conditions.[4]

Proposed Mediating Variables

The following mediating variables were measured on a post-experimental survey: (1) the direction and strength of participants' affective reactions to the auditors and plaintiff; and (2)

[3] I counterbalanced the presentation order of the audit scenarios to control for order effects.

[4] Kadous (2000) notes that it is likely that once a juror has determined that an auditor is liable for a plaintiff's loss, the juror will use a second set of factors to determine the amount of the liability.

participants' assessments of auditor competence. To measure participants' affective reactions, the survey asks participants to indicate the direction and strength of their feelings toward the auditors on a scale from −100 to 100 (−100 = very negative feelings to 100 = very positive feelings). The survey then asks participants to indicate the direction and strength of their feelings toward the plaintiff on the same scale. I subtract participants' responses regarding their feelings for the plaintiff from their responses regarding their feelings for the auditor to create an overall measure of participants' affective reactions to the case (−200 = very strong pro-plaintiff affective reaction; 0 = neutral overall affective reaction; 200 = very strong pro-auditor affective reaction).[5] To measure participants' assessments of auditor competence, the survey asks participants to indicate how competent, on a scale from 0 to 100, they believe the auditors were in their performance of the audit.

IV. RESULTS

The experimental case includes several questions to examine the salience of the manipulations. In particular, the case asks participants about the nature of the auditors' procedures to investigate for the perpetrated fraud. Participant responses to these questions indicate that they attended to and understood the manipulations. Excluding data from participants who did not answer the manipulation checks correctly does not change the inferences drawn from tests of H1, H2, or H3.

The experiment was conducted at different locations on separate dates approximately 22 months apart. The additional participant responses provided by the second data set allow for a more accurate structural-equations based analysis of this study's mediating variables (Lei and Lomax 2005). The inferences drawn from tests of H1, H2, and H3 are the same across both data sets. In addition, there are no significant differences in participant responses for any dependent or mediating variables across the two data sets. Thus, I combine participants' responses from both data sets when reporting the results of tests of H1, H2, H3, and the results of the study's structural equation-based path analysis.

Test of H1

To test H1, I compare the intensity of participants' counterfactual thoughts when the auditors performed audit procedures to investigate for the fraud relative to when the auditors did not investigate for the fraud (see Table 1). Participants' counterfactual thoughts are more intense in the high-investigation condition (Mean (M) = 60.79) than in the no-investigation condition (M = 45.86), ($t_{150} = 3.57$, $p < 0.001$, two-tailed). In addition, participants' counterfactual thoughts are more intense in the low-investigation condition (M = 63.00) than in the no-investigation condition (M = 45.86), ($t_{151} = 4.14$, $p < 0.001$, two-tailed). Therefore, supporting H1, results indicate that evaluators in cases of undetected fraud experience more intense thoughts of what the auditors could have done differently to detect the fraud when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. Furthermore, this result holds across two levels of fraud investigation.[6]

---

[5] Macrae (1992) demonstrates that counterfactual reasoning affects evaluators' affective reactions to both the plaintiff and the defense in civil litigation contexts. Thus, the effect of evaluators' affective reactions on their assessments of auditor liability likely depends on the evaluators' affective reactions to both the auditors and the plaintiff. Specifically, an evaluator's negative feelings toward the auditors will be more likely to increase her assessment of auditor liability if the evaluator has positive feelings toward the plaintiff than if she has negative feelings toward the plaintiff. Similarly, an evaluator's positive feelings toward the plaintiff will be more likely to increase her assessment of auditor liability if the evaluator has negative feelings toward the auditors than if she has positive feelings toward the auditors.

[6] The intensity of participants' counterfactual thoughts did not significantly vary across the low- and high-investigation conditions ($t_{151} = 0.546$, $p = 0.586$, two-tailed).

TABLE 1

Between-Participants Results

Panel A: Descriptive Statistics

Experimental Condition	Counterfactual Intensity Mean (Std. Dev.)	Negligent (Yes/No) Proportion (Percent)	Probability Negligent Mean (Std. Dev.)	Damages Awarded Mean (Std. Dev.)
No investigation (n = 76)	45.86 (26.35)	30/76 (39.47)	40.63 (27.21)	$1.94mil (2.72)
Low investigation (n = 77)	63.00 (24.86)	54/77 (70.13)	51.75 (24.97)	$3.82mil (3.09)
High investigation (n = 76)	60.79 (25.19)	55/76 (72.36)	53.03 (25.06)	$3.56mil (2.69)
Column total	56.58 (26.48)	139/229 (60.70)	48.48 (26.25)	$3.11mil (2.95)

Panel B: Planned Contrasts H1 and H2
No Investigation versus High Investigation

H1: Counterfactual intensity	$t = 3.57$	$p < 0.001$
H2: Negligent (yes/no)	$\chi^2 = 16.68$	$p < 0.001$
H2: Probability Negligent	$t = 2.92$	$p = 0.004$
H2: Damages Awarded	$t = 3.67$	$p < 0.001$

Panel C: Planned Contrasts H1 and H2
No Investigation versus Low Investigation

H1: Counterfactual Intensity	$t = 4.14$	$p < 0.001$
H2: Negligent (yes/no)	$\chi^2 = 14.52$	$p < 0.001$
H2: Probability Negligent	$t = 2.64$	$p = 0.009$
H2: Damages Awarded	$t = 3.92$	$p < 0.001$

Panel A reports the descriptive statistics for participants' responses regarding the intensity of participants' counterfactual thoughts and their between-participants assessments of auditor liability across experimental conditions. The Counterfactual Intensity column reports participants' responses regarding the intensity of their counterfactual thoughts regarding what the auditors could have done differently to detect the fraud. The Negligent (Yes/No) column reports participants' responses as to whether (yes/no) the auditors were negligent in performing the audit. The Probability Negligent column reports participants' probability assessments of whether the auditors were negligent. The Damages Awarded column reports the amount of damages participants' awarded to the plaintiff.

In the no-investigation condition, the auditors did not identify the perpetrated fraud as a fraud risk and did not perform audit procedures to investigate the fraud risk. In the low-investigation condition, the auditors identified the perpetrated fraud as a fraud risk and investigated for the fraud exclusively by making inquiries of the mining company's president. In the high-investigation condition, the auditors investigated for the fraud by making inquiries of the president and also by hiring a geology firm to test a random sample of the mining company's inactive quarries for the presence of hazardous materials that might cause the EPA to require the mining company to use a more expensive restoration method.

Panel B reports the results of the planned comparisons between participants' between-participants responses in the no-investigation versus the high-investigation conditions. The reported p-value remains significant at a Bonferroni adjusted family-wise α of .05.

Panel C reports the results of the planned comparisons between participants' between-participants responses in the no-investigation versus the low-investigation conditions. The reported p-value remains significant at a Bonferroni adjusted family-wise α of .05.

Test of H2

To test H2, I compare participants' between-participants assessments of auditor liability when the auditors performed audit procedures to investigate for the fraud relative to when the auditors did not investigate for the fraud (see Table 1). Participants are more likely to find the auditors negligent in the high-investigation condition (72.36 percent) than in the no-investigation condition (39.47 percent), χ^2 (1, n = 152) = 16.68, p < 0.001. Next, participants' probability assessments of auditor negligence are greater in the high-investigation condition (M = 53.03) than in the no-investigation condition (M = 40.63), ($t_{150} = 2.92$, p = 0.004, two-tailed). Finally, the monetary damages awarded to the plaintiff are greater in the high-investigation condition (M = $3.56 mil) than in the no-investigation condition (M = $1.94 mil), ($t_{148} = 3.67$, p < 0.001, two-tailed).

In addition, participants are more likely to find the auditors negligent in the low-investigation condition (70.13 percent) than in the no-investigation condition (39.47 percent), χ^2 (1, n = 153) = 14.52, p < 0.001. Next, participants' probability assessments of auditor negligence are greater in the low-investigation condition (M = 51.75) than in the no-investigation condition (M = 40.63), ($t_{151} = 2.64$, p = 0.009, two-tailed). Finally, monetary damages awarded to the plaintiff are greater in the low-investigation condition (M = $3.82 mil) than in the no-investigation condition (M = $1.94 mil), ($t_{145} = 3.92$, p < 0.001, two-tailed). Therefore, participants' between-participants assessments of auditor liability are consistently more severe when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. In addition, these results hold across two levels of fraud investigation.[7] These findings support H2.

Structural Equations-Based Path Analysis

The theory underlying this study holds that particular factors mediate the relationship between whether the auditors performed audit procedures to investigate for the fraud and the participants' assessments of auditor liability in cases of undetected fraud. Table 2 reports descriptive statistics for these mediating variables. Table 3 reports simple correlations among the mediators, participants' probability assessments of auditor negligence and the manipulated variable (i.e., whether the auditors investigated for the fraud). To examine the hypothesized relationships between the auditors' decision to investigate for the fraud, the study's mediating variables, and participants' probability assessments of auditor negligence, I use structural equations-based path analysis.

The study's primary construct of interest is evaluators' assessments of auditor negligence, as opposed to evaluators' subsequent assessments of damage amounts. Therefore, I examine participants' assessments of auditor negligence, as opposed to their damage assessments, in the structural equation-based path analysis. I examine participants' probability assessments of auditor negligence, as opposed to participants' yes/no decisions of auditor negligence, because structural equation models generally use continuous variables as opposed to binary categorical variables, as continuous variables are less likely than categorical variables to violate normality assumptions within structural equation models (Kline 1998).[8]

---

[7] Participant assessments of auditor liability did not vary across the low- and high-investigation conditions regarding judgments (yes/no) of auditor negligence χ^2 (1, n = 153) = 0.09, p = 0.759), probability assessments of auditor negligence ($t_{151} = 0.315$, p = 0.753), or damage assessments ($t_{147} = 0.55$, p = 0.583).

[8] It is possible to use categorical variables in standard structural equation models if there is not a severe violation of the normality assumption. A generally accepted rule is that a categorical variable can be used in a standard structural equation model if the distribution's kurtosis does not exceed +/−1.5 (Schumacker and Lomax 2004). The kurtosis of the distribution of participants' yes/no decisions of auditor negligence is −1.82, which is outside of the acceptable range.

TABLE 2

Descriptive Statistics for Mediating Variables

Experimental Condition	Counterfactual Intensity Mean (Std. Dev.)	Participants' Affective Reactions Mean (Std. Dev.)	Participants' Assessments of Auditor Competence Mean (Std. Dev.)
No investigation (n = 76)	45.86 (26.35)	−3.03 (45.54)	68.68 (20.98)
Low investigation (n = 77)	63 (24.86)	−23.34 (48.27)	60.29 (21.45)
High investigation (n = 76)	60.79 (25.19)	−18.11 (46.65)	62.28 (23.07)
Column Total	56.58 (26.48)	−14.86 (47.43)	63.73 (22.05)

This table reports the mean and (standard deviation) of the study's proposed mediating variables by condition. See Table 1 for a description of the experimental conditions.

Descriptions of mediating variables are below.

"Counterfactual Intensity" refers to participants' responses on a scale from 0 to 100 as to how seriously they thought about what the auditors could have done differently to prevent the audit failure.

"Participants' Affective Reactions" refers to the strength of participants' overall affective reactions to the case (−200 = very strong pro-plaintiff reaction, 0 = neutral overall reaction, 200 = very strong pro-auditor reaction).

"Participants' Assessments of Auditor Competence" refers to participants' assessments on a scale from 0 to 100 as to the competence of the auditors.

In performing the structural equations-based path analysis, I combine participants' responses in the high- and low-investigation conditions into one group. Combining the high- and low-investigation conditions is appropriate because the study's theory and hypotheses pertain to whether the auditors performed audit procedures to investigate for the fraud, as opposed to the extent of the procedures performed. In addition, there are no significant differences between the high- and low-investigation conditions regarding the proposed mediating variables or dependent variables. Examining the low- and high-investigation conditions individually does not change the inferences from the path model.

A likelihood ratio Chi-square test indicates the model-implied covariance matrix does not differ from the observed covariance matrix χ^2 (3, n = 229) = 5.01, p = 0.171. The Tucker-Lewis Index indicates the proportion of improvement of the fit of the model over a null model is 0.960 and the incremental fit index is 0.988, both of which are above the standard cutoff value of an acceptable fit of 0.90 (Kline 1998). Thus, the model describes the relationships in the data well. I use t-tests to test the significance of each path coefficient.

Figure 2 presents the path model with unstandardized coefficients for each hypothesized link. Consistent with H1, the relationship between whether the auditors investigated for the fraud and the intensity of participants' counterfactual thoughts of what the auditors could have done differently to detect the fraud (link 1) is significantly positive ($t_{227} = 4.51$, $p < 0.001$, two-tailed). The relationship between the intensity of participants' counterfactual thoughts and their affective reactions (link 2) is significantly negative—i.e., as participants' counterfactual thoughts increased in intensity, their affective reactions were less favorable to the auditor, ($t_{227} = 4.89$, $p < 0.001$,

TABLE 3

Correlation Matrix for Mediating Variables (between-participants environment)

	Auditor Investigation	Counter-factual Intensity	Participants' Affective Reactions	Participants' Assessment of Auditor Competence	Probability Assessment of Auditor Negligence
Auditor investigation	1				
Counterfactual intensity	0.286**	1			
Participants' affective reactions	−0.176**	−0.308**	1		
Participants' assessments of auditor competence	−0.159*	−0.249**	0.338**	1	
Assessment of the probability of auditor negligence	0.211**	0.320**	−0.361**	−0.528**	1

*, ** Indicate the estimated coefficients are significantly different from 0 at 5 percent and 1 percent levels respectively, in two-tailed tests.

This table reports the Pearson correlations between the study's independent variable i.e., whether the auditors investigated for the perpetrated fraud, the study's proposed mediating variables, and participants' probability assessments of auditor negligence.

"Auditor investigation" refers to whether the auditors performed audit procedures to investigate for the perpetrated fraud.

"Counterfactual intensity" refers to participants' responses on a scale from 0 to 100 as to how seriously they thought about what the auditors could have done differently to prevent the audit failure.

"Participants' affective reactions" refers to the strength of participants' overall affective reactions to the case (−200 = very strong pro-plaintiff reaction, 0 = neutral overall reaction, 200 = very strong pro-auditor reaction).

"Participants' assessments of auditor competence" refers to participants' assessments on a scale from 0 to 100 as to the competence of the auditors.

"Assessment of the probability of auditor negligence" refers to participants' assessments of the probability that the auditors were negligent in their performance of the audit.

two-tailed). The relationship between the intensity of participants' counterfactual thoughts and participants' assessments of auditor competence (link 3) is significantly negative ($t_{227} = 2.48$, $p = 0.013$, two-tailed). The relationship between evaluators' affective reactions and their assessments of auditor competence (link 4) is significantly positive ($t_{227} = 4.46$, $p < 0.001$, two-tailed), indicating that participants' assessments of auditor competence increased (decreased) as their affective reactions were more (less) favorable to the auditor.

As expected, the direct relationship between the intensity of participants' counterfactual thoughts and their probability assessments of auditor negligence (link 5) is significantly positive ($t_{227} = 2.82$, $p = 0.005$, two-tailed). The relationship between participants' affective reactions to the case and their assessments of auditor negligence (link 6) is significantly negative ($t_{227} = 2.80$, $p = 0.005$, two-tailed)—i.e., as participants' affective reactions were less favorable to the auditor, their assessments of auditor negligence increased. The relationship between participants' assessments of auditor competence and their assessments of auditor negligence (link 7) is significantly negative ($t_{227} = 7.44$, $p < 0.001$, two-tailed).[9]

[9] A separate model that includes a direct link between whether the auditors investigated for the fraud and participant assessments of auditor liability (along with the links between the proposed mediating variable as found in Figure 2) indicates the direct relationship between whether the auditors investigated for the fraud and participants' probability assessments of auditor negligence is not significant ($t_{227} = 1.33$, $p = 0.183$, two-tailed). However, if one removes the proposed mediating variables from the model, then the direct relationship between whether the auditors investigated for

In summary, consistent with the hypothesized predictions, path analysis indicates that, in cases of undetected fraud, evaluators experience more intense thoughts of what the auditors could have done differently to detect the fraud when the auditors performed audit procedures to investigate for the fraud. As a result, evaluators' affective reactions to the case are less favorable to the auditors and their assessments of auditor competence are lower when the auditors performed audit procedures to investigate for the fraud. In turn, these effects increase evaluators' probability assessments of auditor negligence.

Sensitivity analysis was conducted by calculating modification indices for the model. Modification indices estimate the increase in the overall model's fit by adding specific paths not included in the current model. Thus, modification indices indicate whether model components are connected in a way not specified by the model (Kadous et al. 2005). An examination of modification indices suggests that no additional paths will increase the model's fit, using the conventional 0.05 level of statistical significance.

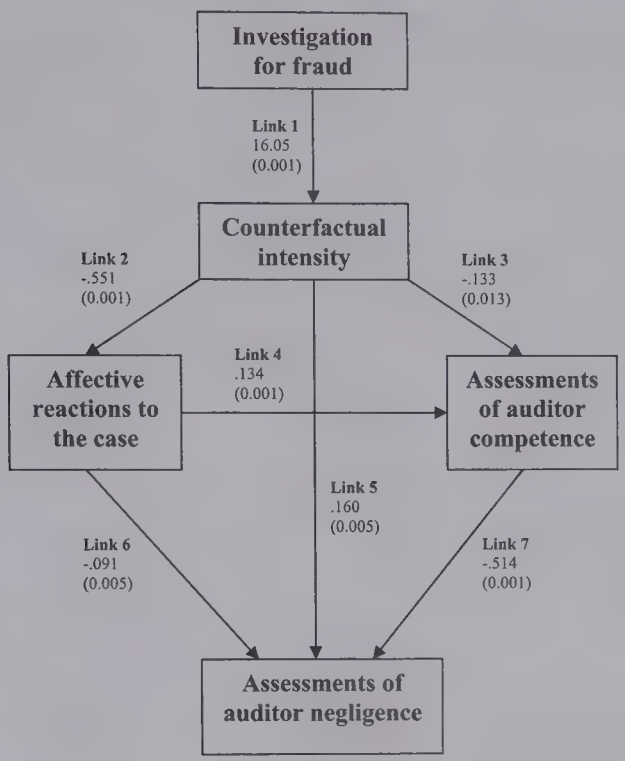
Additional Analysis

Recall from Section II, prior research suggests two potential reasons why evaluators have more intense counterfactual thoughts of what the auditors could have done differently to detect the fraud when the auditors investigated for the fraud. The first reason pertains to outcome closeness (Kahneman and Miller 1986) and the second pertains to the relationship between action (versus inaction) and counterfactual reasoning (Sherman and McConnell 1995). To investigate the effects of evaluators' perceptions of outcome closeness, I measure, on the post-experimental survey, participant perceptions of how close the auditors were to detecting the fraud (0 to 100 scale). Consistent with expectations, participants perceived auditors who investigated for the fraud as having been closer to detecting the fraud than auditors who did not ($t_{227} = 4.86$, $p < 0.001$, two-tailed). In addition, participants' perceptions of outcome closeness are positively correlated with the intensity of their counterfactual thoughts ($p = 0.061$, two-tailed). However, the relationship between participants' perceptions of outcome closeness and the intensity of their counterfactual thoughts is not significant when the analysis controls for whether the auditors investigated for the fraud ($t_{227} = 0.72$, $p = 0.473$, two-tailed). This result indicates that more intense evaluator counterfactual thoughts when the auditors investigated for the fraud (link 1) are not attributable to increased perceptions of outcome closeness, but rather to the fact the auditors took action to investigate for the fraud (Gleicher et al. 1990; Sherman and McConnell 1995). Kahneman and Miller (1986) assert that the intensity of evaluators' counterfactual thoughts increases with the mutability of prior events. The results suggest that the fact that the auditors took action to investigate for the fraud makes the auditors' failure to detect the fraud more mutable. Specifically, results suggest that when evaluators know that the auditors took action to investigate for a perpetrated fraud, it is relatively easy for the evaluators to imagine the auditors adjusting their audit procedures in manner that would have allowed them to detect the fraud. However, when evaluators know that the auditors did not identify the perpetrated fraud as a fraud risk and, thus, did not

---

the fraud and participants' probability assessments of auditor negligence is significantly positive ($t_{227} = 3.27$, $p = 0.001$, two-tailed). These results suggest the effect of investigating fraud risks on participants' assessments of auditor negligence in cases of undetected fraud is mediated by the intensity of participants' counterfactual thoughts, their affective reactions to the auditor and plaintiff, and their assessments of auditor competence. Sobel tests for mediation indicate that each mediating variable is a significant mediator when examined individually. First, the intensity of participants' counterfactual thoughts, when examined individually, mediates the effect of auditor investigation on evaluators' assessments of auditor negligence in cases of undetected fraud ($Z = 3.11$, $p = 0.002$, two-tailed). Next, participants' affective reactions, when examined individually, mediate the effect of auditor investigation on evaluators' assessments of auditor negligence ($Z = 2.42$, $p = 0.016$, two-tailed). Finally, participants' assessments of auditor competence, when examined individually, mediate the effect of auditor investigation on evaluators' assessments of auditor negligence ($Z = 2.34$, $p = 0.019$, two-tailed).

FIGURE 2
Empirical Model of the Effects of Investigating Fraud Risks on Evaluator Assessments of Auditor Liability



This figure shows the results of the path analysis that simultaneously tests the relationships among the proposed mediating variables. Unstandardized path coefficient and corresponding p-value (two-tailed) from t-tests are shown next to each path. Overall goodness of fit is measured through the Tucker-Lewis index (96.0 percent), which measures the proportion of improvement of the fit of the model over a null model, and confirmed with a Chi-square test, χ^2 (3, n = 229) = 5.01, p = 0.171, and an incremental fit index (98.8 percent).
"Investigation for fraud" refers to whether the auditors performed audit procedures to investigate for the fraud. In performing the above path analysis, I combined participants' responses in the high- and low-investigation conditions into one group. Combining the high- and low-investigation conditions is appropriate because the paper's theory and hypotheses pertain to whether the auditors performed audit procedures to investigate for the fraud, as opposed to the extent of the procedures. In addition, there are no significant differences between the high- and low-investigation conditions regarding the proposed mediating variables or dependent variables. Examining the low- and high-investigation conditions individually does not change the inferences from the path model.
"Counterfactual intensity" refers to the intensity of participants' thoughts of what the auditors could have done differently to detect the fraud.
"Affective reactions to the case" refers to participants' affective reactions to the case. Positive (negative) numbers are indicative of affective reactions that are favorable to the auditor (plaintiff).
"Assessments of auditor competence" refers to participants' assessments of how competent the auditors were in their performance of the audit.
"Assessments of auditor negligence" refers to participants' probability assessments of how likely it was that the auditors were negligent in their performance of the audit.

investigate for the fraud, it is relatively difficult for the evaluators to imagine the auditors first identifying the fraud as a fraud risk and then deciding to perform the correct set of audit procedures necessary to detect the fraud.

Test of H3

To test H3, I compare participants' within-participants assessments of auditor liability when the auditors investigated for the fraud relative to when the auditors did not investigate for the fraud (see Table 4). A Chi-square test from a generalized estimating equation model indicates that the probability of participants in a within-participants environment finding the auditors negligent is

TABLE 4

Within-Participants Results

Experimental Condition	Negligent (Yes/No) Proportion (Percent)	Probability Negligent Mean (Std. Dev.)	Damages Awarded Mean (Std. Dev.)
No investigation (n = 225)	187/225 (83.11)	68.25 (27.00)	$5.28mil (3.04)
Low investigation (n = 227)	154/227 (67.84)	55.35 (31.05)	$3.92mil (3.25)
High investigation (n = 226)	75/226 (33.19)	28.03 (30.24)	$1.78mil (2.74)
Column Total	416/678 (61.36)	50.58 (33.88)	$3.65mil (3.34)

Panel B: Planned Contrasts H3
No Investigation versus High Investigation

H3: Negligent (yes/no)	$\chi^2 = 86.86$	$p < 0.001$
H3: Probability negligent	$t = 14.25$	$p < 0.001$
H3: Damages awarded	$t = 12.10$	$p < 0.001$

Panel C: Planned Contrasts H3
No Investigation versus Low Investigation

H3: Negligent (yes/no)	$\chi^2 = 12.53$	$p < 0.001$
H3: Probability negligent	$t = 4.75$	$p < 0.001$
H3: Damages awarded	$t = 4.64$	$p < 0.001$

Panel A reports the descriptive statistics for participants' responses regarding the intensity of participants' counterfactual thoughts and their within-participants assessments of auditor liability across experimental conditions. See Table 1 for a description of experimental conditions and auditor liability assessments.

Panel B reports the results of the planned comparisons between participants' between-participants responses in the no-investigation versus the high-investigation conditions. Chi-square tests are produced by generalized estimating equation and t-tests come from pairwise comparison from a repeated measure ANOVA. The reported p-value remains significant at a Bonferroni adjusted family-wise α of .05.

Panel C reports the results of the planned comparisons between participants' between-participants responses in the no-investigation versus the low-investigation conditions. Chi-square tests are produced by generalized estimating equation and t-tests come from pairwise comparison from a repeated measure ANOVA. The reported p-value remains significant at a Bonferroni adjusted family-wise α of .05.

smaller in the high-investigation condition (33.19 percent) than in the no-investigation condition (83.11 percent) ($\chi^2 = 86.86$, $p < 0.001$).[10] Second, pairwise comparison from a repeated measures ANOVA indicates that participants' within-participants probability assessments of auditor negligence are smaller in the high-investigation condition (M = 28.03) than in the no-investigation condition (M = 68.25), ($t_{262} = 14.25$, $p < 0.001$, two-tailed). Finally, pairwise comparison from a repeated measures ANOVA indicates that participants' within-participants damage awards are smaller in the high-investigation condition (M = \$1.78 mil) than in the no-investigation condition (M = \$5.28 mil), ($t_{260} = 12.10$, $p < 0.001$, two-tailed).

In addition, a Chi-square test from a generalized estimating equation model indicates that the probability of participants in a within-participants environment finding the auditors negligent is smaller in the low-investigation condition (67.84 percent) than in the no-investigation condition (83.11 percent), ($\chi^2 = 12.53$, $p < 0.001$). Second, pairwise comparison from a repeated measures ANOVA indicates that participants' within-participants probability assessments of auditor negligence are smaller in the low-investigation condition (M = 55.35) than in the no-investigation condition (M = 68.25), ($t_{262} = 4.75$, $p < 0.001$, two-tailed). Finally, pairwise comparison from a repeated measures ANOVA indicates that participants' within-participants monetary damage awards are smaller in the low-investigation condition (M = \$3.92 mil) than in the no-investigation condition (M = \$5.28 mil), ($t_{260} = 4.64$, $p < 0.001$, two-tailed). Therefore, participants' within-participants assessments of auditor liability are consistently less severe when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. These results support H3.[11]

V. SUMMARY AND DISCUSSION

Practitioners and legal scholars have expressed concern that the U.S. legal system might, in cases of undetected fraud, effectively penalize auditors for identifying and investigating fraud risks (AICPA 2004; Coffee 2004; Golden et al. 2006). This study provides relevant psychology theory and experimental evidence to indicate whether such concern is warranted. Consistent with counterfactual reasoning theory, evaluators in a natural (i.e., between-participants) setting are *more* likely to hold auditors liable for failing to detect fraud when the auditors performed audit procedures to investigate for the fraud, relative to when the auditors did not investigate for the fraud. Therefore, the study suggests that recent concern regarding the potential legal hazards of investigating fraud risks is warranted. While I do not examine auditors' overall incentives to investigate fraud risks, the possibility of the legal system, in any circumstance, effectively penalizing auditors for investigating fraud risks is troubling. Future research is therefore needed to further assess the severity of this finding.

In addition, I find that evaluators in a within-participants setting are *less* likely to hold auditors liable for failing to detect fraud when the auditors performed audit procedures to investigate for the fraud. This finding demonstrates that evaluators understand that higher levels of

---

[10] Generalized estimating equations permit specification of a working correlation matrix that controls for within-subject correlation of responses on dependent variables of several different distributions, including normal, binomial, and Poisson. Consequently, generalized estimating equations estimate more efficient and unbiased regression parameters relative to ordinary least-squares regression, particularly for dichotomous dependent variables (Ballinger 2004).

[11] Within-participant assessments of auditor liability are consistently less severe in the high- versus the low-investigation condition. A Chi-square test from a generalized estimating equation model indicates that the probability of participants in a within-participants environment finding the auditors negligent is smaller in the high- versus the low-investigation condition ($\chi^2 = 63.20$, $p < 0.001$). Second, pairwise comparison from a repeated measures ANOVA indicates that participants' within-participants probability assessments of auditor negligence are smaller in the high- versus the low-investigation condition ($t_{262} = 10.93$, $p < 0.001$). Finally, pairwise comparison from a repeated measures ANOVA indicates that participants' within-participants damage assessments are smaller in the high- versus the low-investigation condition ($t_{260} = 8.61$, $p < 0.001$).

auditor investigation imply lower levels of auditor liability and that evaluators intend to apply this standard in cases of undetected fraud. Consequently, the observed between-participants decision pattern appears to be unintentional (Kahneman and Tversky 1996). This is important because it suggests that a debiasing method could potentially reverse the troubling between-participants decision pattern (Wegener et al. 2000). Prior research identifies several methods of debiasing evaluators' assessments of auditor liability (e.g., Anderson et al. 1997; Kadous 2001; Clarkson et al. 2002; Grenier et al. 2009). The knowledge gained from prior studies and from this study provides direction for future research attempting to identify an effective method of reversing evaluators' unintentional tendency, when evaluating individual cases of undetected fraud, to judge auditors who investigated for the fraud more harshly than auditors who did not investigate for fraud.

This study also extends psychology research examining the effects of counterfactual reasoning on evaluators' assessments of defendant liability in litigation settings. Prior studies have examined settings where the level of care provided by the defendant is constant across high- and low-counterfactual conditions. Conversely, the level of care provided by the defendant (i.e., the auditor) in this study varies between conditions. I find that the intensity of evaluators' counterfactual thoughts increases with the auditors' level of fraud investigation and, thus, also increases with the level of care provided by the auditor. In addition, consistent with prior research, I find that the severity of evaluators' assessments of auditor liability increases with the intensity of their counterfactual thoughts. These findings suggest that the effects of counterfactual reasoning increase evaluators' assessments of defendant liability in settings where more intense counterfactual reasoning is associated with higher levels of defendant care.

This study is subject to limitations. First, consistent with cases used in prior research, differences in the auditors' actions across conditions require that the plaintiff and defense attorneys' arguments vary between conditions. In the no-investigation condition, the plaintiff attorneys argue that the lack of awareness and consequent investigation of a major fraud risk indicates auditor negligence. In the low- and high-investigation conditions, the plaintiff attorneys argue that the failure to detect a fraud that the auditors investigated indicates the auditors did not effectively investigate a known fraud risk and were therefore negligent. In all conditions, the defense attorneys argue the fraud is extremely rare, that other auditors would not have detected the fraud, and that auditors cannot be expected to detect all frauds. In addition, in the high-investigation condition, the defense attorney compels the plaintiff's expert witness to concede that sample testing is consistent with generally accepted auditing standards. The differences across conditions raise the possibility that the between-participants results could be either partially or fully attributable to differences in the strength of the attorneys' arguments across conditions as opposed to the presented theory. There are several reasons, however, why I believe this possibility is unlikely.

First, prior to running the experiment, a legal scholar reviewed the experimental materials for realism and ensured the consistency of the strength of the wording of the attorneys' arguments across conditions. I incorporated her suggestions into the experimental materials to her satisfaction that the experimental case was realistic and that the strength of the attorneys' language did not vary across conditions. Second, several reported results collectively suggest the observed between-participants effects are attributable to the presented theory, as opposed to differences in the strength of the attorneys' arguments. Specifically, there is no significant difference in the intensity of participants' counterfactual thoughts across the low- and high-investigation conditions. Yet, the content of the attorneys' arguments significantly varied across these conditions, which suggests that the intensity of evaluators' counterfactual thoughts was not a function of the content of the attorneys' arguments but, rather, a function of whether the auditors performed audit procedures to

investigate for the fraud. Also, within each experimental condition (and thus holding the content of the attorneys' arguments constant), there is a significant positive correlation between the intensity of participants' counterfactual thoughts and their assessments of auditor negligence. This indicates that more intense counterfactual reasoning was associated with harsher assessments of auditor liability, irrespective of the content of the attorneys' arguments.

Finally, it is important to note that the expected litigation cost associated with an audit is a function of two factors: (1) the probability of audit failure, and (2) the expected litigation cost given audit failure. This study examines the expected litigation costs given audit failure, but not the probability of audit failure. This study is, therefore, silent as to the overall effect of investigating fraud risks on auditors' expected litigation costs. Accordingly, this study does not imply that auditors should avoid investigating fraud risks.

REFERENCES

Alicke, M., J. Buckingham, E. Zell, and T. Davis. 2008. Culpable control and counterfactual reasoning in the psychology of blame. *Personality and Social Psychology Bulletin* 34 (10): 1371–1381.

American Institute of Certified Public Accountants (AICPA). 2002. *Consideration of Fraud in a Financial Statement Audit*. Statement on Auditing Standards No. 99. New York, NY: AICPA.

———. 2004. *AICPA Practice Aid Series, Fraud Detection in a GAAS Audit: SAS No. 99 Implementation Guide*. Revised edition. New York, NY: AICPA.

———. 2006. *Risk Assessment Standards*. Statement on Auditing Standards Nos. 104–111. New York, NY: AICPA.

Anderson, J. C., M. M. Jennings, D. J. Lowe, and P. J. Reckers. 1997. The mitigation of hindsight bias in judges' evaluation of auditor decisions. *Auditing: A Journal of Practice & Theory* 16 (2): 20–39.

Ballinger, G. A. 2004. Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods* 7: 127–150.

Bornstein, B. H. 1999. The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior* 23: 75–91.

Bothwell, R., and K. Duhon. 1994. Counterfactual thinking and victim compensation. *The Journal of Social Psychology* 134 (5): 705–706.

Bower, G. H. 1981. Mood and memory. *The American Psychologist* 36: 129–148.

Bower, G., and J. Forgas. 2001. Mood and social memory. In *Handbook of Affect and Social Cognition*, edited by Forgas, J., 95–110. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Causey, D., and S. A. Causey. 1991. *Duties and Liabilities of Public Accountants*. 4th edition. Mississippi State, MS: Accountant's Press.

Clarkson, P., C. Emby, and V. Watt. 2002. Debiasing the outcome effect: The role of instructions in an audit litigation setting. *Auditing* 21 (2): 7–20.

Coffee, J. 2004. Gatekeeper failure and reform: The challenge of fashioning relevant reform. *Boston University Law Review. Boston University. School of Law* 84: 301–332.

Cornell, R., R. Warne, and M. Eining. 2009. The use of remedial tactics in negligence litigation. *Contemporary Accounting Research* 26 (3): 767–787.

Creyer, E. H., and Z. Gurhan. 1997. Who's to blame? Counterfactual reasoning and the assignment of blame. *Psychology and Marketing* 14: 209–222.

Frank, K. E., D. J. Lowe, and J. Smith. 2001. The expectation gap: Perceptual differences between auditors, jurors and students. *Managerial Auditing Journal* 3: 45–150.

Gleicher, E., K. A. Kost, S. M. Baker, A. J. Strathman, S. A. Richman, and S. Sherman. 1990. The role of counterfactual thinking in judgments of affect. *Personality and Social Psychology Bulletin* 16: 284–295.

Golden, T., S. Skalak, and M. Clayton. 2006. *A Guide to Forensic Accounting Investigation*. Hoboken, NJ: John Wiley and Sons.

Grenier, J., M. Peecher, and D. Piercy. 2009. Judging auditor negligence: Debiasing interventions, outcome bias, and reverse outcome bias. Working paper, University of Illinois Urbana–Champaign.

Kadous, K. 2000. The effects of audit quality and consequence severity on juror evaluations of auditor responsibility for plaintiff losses. *The Accounting Review* 75 (3): 327–341.

———. 2001. Improving jurors' evaluations of auditors in negligence cases. *Contemporary Accounting Research* 18 (3): 425–444.

———, L. Koonce, and K. Towry. 2005. Quantification and persuasion in managerial judgment. *Contemporary Accounting Research* 22 (3): 643–686.

Kahneman, D., and D. T. Miller. 1986. Norm theory: Comparing reality to its alternatives. *Psychological Review* 93: 136–153.

———, and A. Tversky. 1996. On the reality of cognitive illusions. *Psychological Review* 103 (3): 582–588.

Kline, R. B. 1998. *Principles and Practices of Structural Equation Modeling*. New York, NY: Guilford Press.

Landman, J. 1987. Regret and elation following action and inaction: Affective responses to positive versus negative outcomes. *Personality and Social Psychology Bulletin* 13: 524–536.

Lei, M., and R. Lomax. 2005. The effects of varying degrees of nonnormality in structural equation modeling. *Structural Equation Modeling* 12: 1–27.

Lowe, D. J., P. Reckers, and S. Whitecotton. 2002. The effects of decision-aid use and reliability on jurors' evaluations of auditor liability. *The Accounting Review* 77 (1): 185–202.

Macrae, C. N. 1992. A tale of two curries: Counterfactual thinking and accident-related judgments. *Personality and Social Psychology Bulletin* 18: 84–87.

Markman, K. D., I. Gavinski, S. Sherman, and M. McMullen. 1993. The mental simulation of better and worse worlds. *Journal of Experimental Social Psychology* 29: 87–109.

McMullen, M. N., K. D. Markman, and I. Gavanski. 1995. Living in neither the best nor the worst of all possible worlds: Antecedents and consequences of upward and downward counterfactual thinking. In *What Might Have Been: The Social Psychology of Counterfactual Thinking*, edited by Roese, N., and J. Olson, 133–167. Mahwah, NJ: Erlbaum.

Medvec, V. H., S. E. Madey, and T. Gilovich. 1995. When less is more: Counterfactual thinking and satisfaction among olympic athletes. *Journal of Personality and Social Psychology* 69: 603–610.

Miller, D. T., and C. McFarland. 1986. Counterfactual thinking and victim compensation: A test of Norm Theory. *Personality and Social Psychology Bulletin* 12: 513–519.

Peecher, M., and D. Piercey. 2008. Judging audit quality in light of adverse outcomes: Evidence of outcome bias and reverse outcome bias. *Contemporary Accounting Research* 25 (1): 243–274.

Radhakrishnan, S. 1999. Investors' recovery friction and auditor liability rules. *The Accounting Review* 74: 225–240.

Robbennolt, J. K., J. Darley, and R. MacCoun. 2003. Symbolism and incommensurability in civil sanctioning: Legal decision-makers as goal managers. *Brooklyn Law Review* 68: 1121–1158.

———, and M. S. Sobus. 1997. An integration of hindsight bias and counterfactual thinking: Decision-making and drug courier profiles. *Law and Human Behavior* 21 (5): 539–560.

Roese, N. 1994. The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology* 65: 805–818.

———. 1997. Counterfactual thinking. *Psychological Bulletin* 121 (1): 133–148.

———, and J. Olson. 1995. *What Might Have Been: The Social Psychology of Counterfactual Thinking*. Mahwah, NJ: Erlbaum.

Sanna, J., and K. Turley-Ames. 2000. Counterfactual intensity. *European Journal of Social Psychology* 30: 273–296.

Schwarz, N., and G. L. Clore. 1983. Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology* 45 (3): 513–523.

Schumacker, R., and R. Lomax. 2004. *A Beginner's Guide to Structural Equation Modeling*. Mahwah, New Jersey: Erlbaum.

Sherman, S. J., and A. R. McConnell. 1995. Dysfunctional implications of counterfactual thinking: When alternatives to reality fail us. In *What Might Have Been: The Social Psychology of Counterfactual Thinking*, edited by Roese, N., and J. Olson, 199–231. Mahwah, NJ: Erlbaum.

Smith, R., and D. Tidrick. 1998. The effect of alternative judicial systems and settlement on auditing. *Review of Accounting Studies* 2: 353–381.

Wegener, D., N. Kerr, N. Flemming, and R. Petty. 2000. Flexible corrections of juror judgments: Implications for jury instructions. *Psychology, Public Policy, and Law* 6 (3): 629–654.

Zickafoose, D., and B. Bornstein. 1999. Double discounting: The effects of comparative negligence on mock juror decision making. *Law and Human Behavior* 23 (5): 577–596.

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 2169–2171

American Accounting Association
DOI: 10.2308/accr.00000004

BOOK REVIEWS

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Business, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

Developments in New Reporting Models, Information for Better Markets (London, U.K.: ICAEW Financial Reporting Faculty, 2009, ISBN 978-84152-615-7, electronic version downloadable from ICAEW.com website, pp. vii, 98).

Developments in New Reporting Models is a discourse on the process of evolution of business reporting models. A prior publication also produced by the Institute of Chartered Accountants of England and Wales (ICAEW), *New Reporting Models for Business* (2003), assessed the merits of a number of proposed business reporting models. The current publication revisits the proposed models briefly and updates the list (Appendix 2), but its objective is not to evaluate specific models. Instead, the report focuses on the mechanisms of change in business reporting models—the factors that facilitate the process and the factors that hinder it. The principal authors of the report are Brian Singleton-Green and Robert Hodgkinson, who are the senior staff of the ICAEW's Information in Betters Markets thought leadership program.

The central premise of the text is that business reporting models evolve naturally over time as they adapt to changes in environmental factors. The authors conclude that, although calls for revolutionary change in business reporting models are a necessary element of the process of evolution, such calls are unlikely to accomplish their goal of widespread adoption of ostensibly "new" business reporting models.

The authors submit that business reporting is shaped by its environment, which includes a society's regulatory, legal, and market setting, and that business reporting models respond to changes in that environment. Nonetheless, the authors recognize that certain frictions can hinder the natural evolution of the business reporting model. Imprudent regulation, excessive litigation risk, the information needs of regulators supplanting those of users, and frictions in the demand and supply of information can interfere with the proper development of business reporting. They suggest that understanding and addressing the root causes of inadequate business reporting is an important, but often overlooked, element of the process of evolution.

Despite being written by non-academics, the authors are markedly committed to the position that less rhetoric and more reliance on research findings would greatly enhance the debate surrounding the adequacy of business reporting models. This commitment is evident in several respects. First, the authors' discussion incorporates relevant findings from research previously commissioned by the ICAEW. Second, throughout the text, the authors identify areas in which they believe additional research would advance the debate. Appendix 5 of the manuscript summarizes the opportunities for further research which the authors identify.

The bulk of the report is organized as follows. Chapter 1 presents the authors' key conclusions that business reporting evolves over time and is a product of its environment. They also introduce the primary charges leveled at existing business reporting models, and in subsequent chapters they evaluate each charge. The authors consider the charges that business reporting is: (a) fundamentally misleading (Chapter 2); (b) provides insufficient reporting of intangibles (Chapter 3); (c) a factor in causing or compounding the recent financial crisis (Chapter 4); (d) offers insufficient reporting of non-financial information (Chapter 5), and; (e) is excessive in length and complexity (Chapter 7).

In each chapter the authors outline the arguments and research evidence (if any) for and against the given charge. The authors do not attempt nor pretend to conduct a comprehensive review of the relevant literature, but

rely primarily on the findings of research previously commissioned by the ICAEW. In addition, they identify areas that they believe are ripe for additional research. In general, the authors conclude that none of the charges leveled at financial reporting is sufficient to yield revolutionary change in existing business reporting models.

Chapter 6 further delves into the development of business reporting, with particular emphasis on the point that business reporting evolves in response to changes within a complex social environment in which business reporting is but one element of a larger informational network. The authors identify factors that might influence the evolution of business reporting in a favorable manner, such as peer pressure and users' demands for information. Nevertheless, the authors note that, if business reporting is indeed broken or breaks in certain respects, it is important to understand and address the factors that interfere with the appropriate evolution of business reporting, and not simply call for dramatic change in the model.

The final chapter of the manuscript outlines a way forward. The authors warn against excessive rhetoric, which they argue can be a distraction. The authors note that rhetoric is often conflicting (e.g., more disclosure is needed but financial statements need to be shorter), and they appeal to those presenting proposed alternative business reporting models to be more firmly grounded in extant research. While I applaud and agree with the authors' plea regarding research, it is important to recognize that various studies' research findings can also be conflicting. Accordingly, it is often necessary to draw conclusions based on the overwhelming weight of the evidence as opposed to relying on the specific findings of a particular study or small group of studies. Moreover, I believe that the authors' plea highlights the importance of collaboration between practitioners, standard setters, and academics.

The authors also consider the role of business reporting within the whole of the information set. They make the point that business reporting is not the only game in town. Other sources of information exist. Moreover, managers suffer from limited information and biases, such that managers might not have a competitive advantage in providing certain types of information. Hence, there is no need for a comprehensive business reporting model, nor would such a model be desirable.

Although the text is quite thorough in its consideration of the issues, at the conclusion of my reading the report, I was left with several questions. First, the authors make the important point that different cultures, rules of law, and legal environments are likely to produce different approaches to business reporting. They do not, however, discuss the implications of this statement for the feasibility of the globalization of accounting standards specifically, which comprise an important component of the business reporting model.

Second, the authors note that the boundaries of "business reporting" and the term "business reporting model" are not well defined, and that the lack of consensus around these terms serves to complicate the debate surrounding the adequacy of business reporting models. While attempting to clarify these concepts is outside the scope of the report, the discussion is subject to a certain amount of slippage with respect to what aspects of reporting are under consideration at various points in the text.

Third, the authors discuss the role of "the invisible hand" in guiding the evolution of business reporting, and note that in certain circumstances regulation might play a beneficial role in overcoming demand or supply frictions that could hinder adequate business reporting. They do not, however, discuss the extent to which regulatory bodies do or do not exist locally or internationally to take on the task of developing standards for business reporting beyond the financial statements and footnotes. Perhaps prior calls for revolutionary change in business reporting have failed in part because no obvious entity exists to take up the cause of implementing the proposed changes.

Finally, the authors discuss the extent to which financial statements can or should be "general purpose." They adopt investors as the primary users of business reporting. I believe that making such a choice is necessary. It is difficult to consider whether the information produced by a business reporting model is useful, without first taking a position on the question: useful to whom? While some might disagree with the authors' decision to focus on investors, I find no reason to quibble with this decision. Nonetheless, having made this choice, the authors, I believe, could have used it to provide additional structure to their discussions in Chapters 2 through 8. For example, if investors' needs consist of understanding the past, predicting the future, and assessing risk, the authors' discussion of the charges leveled against existing models could be couched in terms of the extent to which users' needs are, or are not, met. The chapter on intangibles (Chapter 3), in which the authors discuss the claim that the reporting of intangibles on the balance sheet is insufficient, is an example of where such a framework might have provided additional useful structure to the discussion.

In conclusion, I enjoyed reading this thought-provoking report. I commend the authors for contributing a new, and substantive, twist to the debate on business reporting models, by digressing from the established path of assessing the merits of specific proposed models and instead embarking on a fresh avenue of considering the process of change.

REFERENCES

New Reporting Models for Business. 2003. London, U.K.: Institute of Chartered Accountants in England and Wales.

CHRISTINE A. BOTOSAN
Chair in Ethical Financial Reporting
University of Utah

THE ACCOUNTING REVIEW
Vol. 85, No. 6
2010
pp. 2173–2203

American Accounting Association
DOI: 10.2308/accr.00000003

Annual Report and Editorial Commentary for *The Accounting Review*

Steven J. Kachelmeier
The University of Texas at Austin

I. INTRODUCTION

I am pleased to submit the second Annual Report for *The Accounting Review* under the new content guidelines adopted by the Publications Committee and Executive Committee of the American Accounting Association (AAA). From the feedback I have heard, the new format has succeeded in creating a more open environment of full disclosure and accountability to our constituents. It is also a learning experience for the senior editor, providing useful feedback on our processes and decisions. As I did last year, I am taking the liberty of interjecting a liberal dose of editorial commentary along with the statistics, subject to the caveat that my interpretations are opinions with which others may disagree.

Section II below updates last year's qualitative commentary on the editorial process followed by *The Accounting Review* under the current regime. Section III then follows with the tabular data requested by the AAA Publications Committee, along with some supplemental data I have provided to clarify certain aspects of these tables. Section IV closes with some personal notes of thanks and remembrances.

II. UPDATES TO *THE ACCOUNTING REVIEW* EDITORIAL PROCESS

To avoid repetition, I refer the reader to last year's report (Kachelmeier 2009) for a description of the editorial process followed by *The Accounting Review*. This year's description is limited to an update of new developments and changes in process that have occurred over the fiscal year ending May 31, 2010.

One important update is that several coeditors of *The Accounting Review* have generously agreed to increase their allocation beyond the maximum of three new submissions per month that I promised upon establishing our editorial team. Especially for financial-archival manuscripts that constitute so many of our submissions (as reported later), the coeditors who agreed to take on a bit more generally now handle four and sometimes five new submissions per month. As a consequence, my proportion of the total decision letters has declined from last year, in which I wrote 345 (48.0 percent) of the 719 decision letters sent from June 1, 2008 to May 31, 2009. For the fiscal year of this report ending May 31, 2010, I wrote 205 (30.5 percent) of the journal's 673 decision letters, delegating slightly more than two-thirds of our decisions to coeditors whose expertise aligns more closely with the area of the submission. Another reason for the decline in my

I am grateful for helpful feedback and suggestions from Linda Bamber, Jean Bedard, and Tracie Majors on an earlier version of this report, though I assume full responsibility. I also thank Mary Capps for painstakingly checking the current affiliations in the Appendix of *ad hoc* reviewers.

Published Online: November 2010

share of the decision letters is that I wrote most of the decision letters for "transition" manuscripts that were in process at the time the current editorial term began on June 1, 2008. Almost all of those manuscripts have now cleared the system, so current decisions, whether for new submissions or for revisions, tend to be for manuscripts initially submitted under the current regime and assigned to one of the current editors from the onset.

While my decision-letter volume has subsided somewhat, it remains my preference to centralize the reviewer selection process. With the help of my truly outstanding doctoral research assistant, Tracie Majors, I review each new submission and Tracie's search results from various databases to find two well-qualified and independent reviewers, proposing those names to the assigned editor (if other than me). A centralized reviewer selection process helps to ensure uniformity of reviewer credentials across submissions, in addition to facilitating the coordination problem of avoiding multiple requests to the same reviewer at the same time (although this is sometimes unavoidable for invited revisions). Given the demands on our most frequently requested reviewers, a typical strategy is to choose one reviewer from *TAR*'s Editorial Advisory and Review Board (or a substitute of similar seniority) and one other reviewer, likely more junior, whom we ask less often. Other times we try to diversify, such as choosing a reviewer from each of two areas relevant to the submission under consideration. To avoid sending signals, we use a random algorithm to ensure that the labeling of "Reviewer A" and "Reviewer B" is completely arbitrary.

Over the year, we found ourselves asking some *ad hoc* reviewers about as many times as we would ask an Editorial Board member, so in fairness, I generally invited those individuals to join the Board if they had submitted several high-quality reviews on a timely basis. Thus, *TAR*'s Editorial Advisory and Review Board has grown from 118 members at the time of our first issue (January 2009) to 129 members as of the date of this report. An Editorial Board of 129 members might seem large (as recently as 2005, there were only 68 members), but I am continually amazed at how difficult it is to find "open" Editorial Board members who do not already have a review assignment in hand. Sometimes "Professor X" is the obvious choice to review a new submission closely related to X's expertise, but if we recently asked Professor X to review a different manuscript, we generally look elsewhere. In short, we do the best we can to optimize reviewer selection, considering both relevance and availability.

Even with 129 Editorial Board members, it would be highly misleading to infer that these 129 experts write all the reviews. For the fiscal year ending May 31, 2010, *TAR* asked an additional 453 *ad hoc* reviewers, as named and thanked in the Appendix, to evaluate one or more submissions. Thus, the journal's editorial decisions during the fiscal year ending May 31, 2010 were guided by 582 (= 453 + 129) different experts across a wide variety of topical and methodological interests. The diversity inherent in 582 different reviewers is consistent with *TAR*'s mission.

Given the reviewer demands of a journal with over 500 new submissions per year (not even counting revisions), several people have asked me why *TAR* persists with using two different reviewers for each new submission, especially given that some of our competitors use only one. To be sure, a two-reviewer system is costly, both in reviewer resources and in time (by construction, we are always waiting on the later of the two reviewers). *Ex post*, authors often sense that it is more difficult to address the concerns of two reviewers than it would be to address a single reviewer's concerns, especially if they could pick which reviewer to keep and which to discard. But therein lies the point. Editors have to select reviewers *ex ante*, and given that reviewers often disagree (Blank 1991; Gilliland and Cortina 1997; Lynch 1998), both the editor and author gain some protection by getting two draws from the distribution instead of just one. Thus, though authors sometimes see it differently, I honestly see a two-reviewer system as being in the author's best interest, at least *ex ante*. When two reviewers forward different recommendations, editorial judgment becomes critical. The general decision model employed by the current editorial team for

split reviews is to consider whether the negative reviewer has identified a "fatal flaw" that is inherent to the study. If so, we reject, but otherwise we generally move forward, even if that means overruling a reviewer who recommends rejection. Along that line, another benefit of a two-reviewer system is that it provides useful feedback to the reviewers, especially those newer to the process. That is, each reviewer can calibrate his/her assessment against that of the other reviewer, as we send the reviewers copies of both reviews along with a blind copy of the editorial decision letter.

A new development this year is that I have started to enforce the expectation that revisions should be submitted within a year of the decision letter inviting the revision. While it has long been part of *TAR*'s published Editorial Policy, it is my understanding that the "one-year rule" was rarely enforced prior to 2009. Beginning June 1, 2008, however, we started noting this policy explicitly in decision letters inviting a revision. Thus, when the one-year anniversary arrives with no revision in hand, I generally write an email to the submitting author offering one more month (essentially, a grace period) to revise and resubmit. Six times during the fiscal year ending May 31, 2010, after the 13th month elapsed with no revision, I informed the author of my intent to close the file. This might seem harsh, but if we are serious about lessening the time it takes from an initial submission to the ultimate decision, part of that responsibility lies with authors. Moreover, it is my experience that very old revisions fare worse in the review process, as they face the risk of obsolescence (research tends to have a short "shelf life").

There are two caveats to my willingness to enforce the policy allowing one year to revise and resubmit. First, sometimes an author will reply back to my "one-year anniversary" email with an explanation of extenuating medical or personal circumstances that prevented a more timely revision, and I try to take those circumstances into account in reaching a reasonable accommodation to which the author and I can agree. Second, I have not enforced the one-year deadline for revisions invited before the current editorial term began on June 1, 2008, as earlier decision letters did not state that deadline explicitly. By now, however, I hope that the vast majority of those revisions are behind us.

For several other aspects of *TAR*'s editorial process, I encourage the reader to see my previous year's report (Kachelmeier 2009).[1] It is my general understanding from the AAA Publications Committee that each new senior editor will begin his/her term with a report that provides a more comprehensive description of his/her editorial process and philosophy, as I tried to do last year, followed by updates in subsequent years of the editorial term, but without repeating points that have not changed.

III. EDITORIAL AND PUBLICATION STATISTICS

This section of the report provides specific tables requested by the AAA Publications Committee, along with supplemental data for clarification. To differentiate the supplemental material, each table first reports the specific data requested by the Publications Committee, followed by any supplemental data explained in the discussion of each table.

Table 1: Annual Activity Summary

Table 1, Panel A reconciles *TAR*'s workflow for the prior and current journal year ending May 31, 2009 and 2010, respectively. The volume of new submissions is down somewhat, from 557 last year to 502 in fiscal 2010. My best guess is that this difference does not reflect a reversal of *TAR*'s general trend of increasing submissions, but rather reflects a "blip" of an exceptionally large

[1] In addition to last year's annual report, for my personal thoughts on the review process in accounting journals, see Kachelmeier (2004).

TABLE 1
Annual Activity Summary

Panel A: Annual Activity Summary by Journal Year

Journal Year Ending May 31	Manuscripts in Process at Beginning of Year (a)	New Submissions Received (b)	Resubmissions Received (c)	Manuscripts Available for Evaluation (d) = (a) + (b) + (c)	Decision Letters Sent (e)	Manuscripts in Process at End of Year (f) = (d) − (e)
2009	133	557	163	853	719	134
2010	134	502	212	848	673	175

Panel B: New Submissions by Calendar Year

Calendar Year	New Submissions
1998	196
1999	239
2000	260
2001	328
2002	324
2003	327
2004	307
2005	389
2006	413
2007	443
2008	482
2009	508

Explanations of columns in Panel A:

[a] Manuscripts in process include all new submissions and revisions pending decision (generally awaiting review), but exclude manuscripts awaiting revision from authors.

[b] New manuscripts, excluding resubmissions of revised manuscripts.

[c] Resubmissions of invited revisions.

[d] Summation of columns (a), (b), and (c).

[e] Completed decision letters, including subsequent decisions for manuscripts previously evaluated as a "revise and resubmit" decision within the same fiscal year.

[f] Manuscripts awaiting review and/or editorial decision as of the end of the journal's fiscal year.

number of new submissions in the summer months of 2008 from authors who did not want their manuscripts to overlap two editorial regimes. This conjecture is supported by the fact that *TAR* received only 163 new submissions during the first five months of calendar 2008 (i.e., the last five months of former Senior Editor Dan Dhaliwal's term), as compared to 238 new submissions during the first five months of 2009 and 232 new submissions during the first five months of 2010. I suspect that we may experience a similar downturn in spring of 2011 for authors waiting for the new regime to begin on June 1, 2011. Authors should be aware, however, that editorial transition issues are not as complex as they might appear. The current regime will continue to choose reviewers through May 31, 2011, but the next regime will likely be making most of the decisions for manuscripts submitted just before the end of the fiscal year, after the reviews arrive sometime in June or July.

As was the case last year, a limitation of Table 1, Panel A is that logging all submissions results in double counting some manuscript files, to the extent that *TAR* often receives two or more versions of the same manuscript within the same year, following invitations to revise and resubmit. For the journal year ending May 31, 2010 (2009), the 673 (719) total decisions reflect 615 (646) unique manuscript files, with the differences representing revisions of manuscripts for which previous decisions inviting revision had already been logged within the same fiscal year.

Table 1, Panel B is supplemental, reporting submissions by calendar year instead of the journal's fiscal year that ends on May 31. The calendar year tallies were reported prior to 2008, so Panel B facilitates comparison over time. It corroborates that, notwithstanding the temporary surge in new submissions that occurred shortly after June 1, 2008, the general trend of submissions has been steadily increasing for several years. Specifically, over the past decade, *TAR*'s submission volume has approximately doubled.

Table 2: Annual Outcome Summary

Table 2, Panel A: Outcomes by Fiscal Year

Table 2, Panel A tallies the outcomes for all decisions reached during the fiscal year. As requested by the Publications Committee, this panel also reports two "acceptance rates." The first rate in column (e) divides the number of acceptances or conditional acceptances by the number of final accept or reject decisions. The second rate in column (f) divides the same (conditional) acceptances by the total number of decisions reached. The difference between the two denominators reflects invitations to revise and resubmit. Accordingly, the first rate somewhat overstates and the second rate somewhat understates the "true" acceptance rate. As in last year's report, Table 2 includes "conditional" acceptances in the acceptance category, as we have yet to reject any manuscript that has reached "conditional acceptance" status.

As with Table 1, a limitation of Table 2, Panel A is that the total number of decisions in any given journal year includes some manuscript files for which multiple decisions are made within the same year, due to invitations to revise and resubmit. If one divides the number of acceptances by the number of unique manuscript files processed (646 and 615 for the years ending May 31, 2009 and 2010, respectively), the resulting estimated acceptance rates are 12.5 percent (=81/646) for fiscal 2009 and 10.2 percent (=63/615) for fiscal 2010, bearing in mind that the denominator for both years includes several open files awaiting further revision, for which final decisions have yet to be reached.

The reader might have cause for concern in noting that our total number of acceptances is down from 81 last year to 63 this year. The estimated acceptance rates also reflect this downturn. However, I do not perceive any substantive change in our editorial standards for acceptance. One reason for the decline reflects my attempt to build a modest "buffer" of forthcoming articles by taking special efforts to expedite some near-acceptance manuscripts shortly after the current editorial term began on June 1, 2008. Before that time, *TAR* was essentially operating on an issue-

TABLE 2

Annual Outcome Summary

Panel A: Outcomes by Fiscal Year

Journal Year Ending May 31,	Decision Letters Sent (a)	Rejections (b)	Revise and "Uncertain" Decisions (c)	Acceptances and Conditional Acceptances (d)	Acceptance Rate 1: (e) = (d)/[(b) + (d)]	Acceptance Rate 2: (f) = (d)/(a)
2009	719	408	230	81	16.6%	11.3%
2010	673	403	207	63	13.5%	9.4%

Panel B: Final Outcome Resolution for All New Submissions

Journal Year Ending May 31,	New Submissions Received (a)	Acceptances through May 31, 2010 (b)	Rejections through May 31, 2010 (c)	Files Pending Further Revision as of May 31, 2010 (d)	Lower Bound on Acceptance Rate: (e) = (b)/(a)	Upper Bound on Acceptance Rate: (f) = [(b) + (d)]/(a)
2009	557	45	442	70	8.1%	20.6%

Explanations of columns in Panel A:

[a] This column is the same as column (e) of Table 1, Panel A, reflecting all decision letters sent during the fiscal year, including decisions on manuscripts that had already been evaluated previously within the same fiscal year (with invitation to revise and resubmit). Thus, the number of unique manuscript files processed is somewhat lower than the number of decision letters sent.

[b] Exhibit 1 separates the total rejections into manuscripts rejected due primarily to insufficient contribution and manuscripts rejected due primarily to a perceived threat to the validity of the reported claims.

[c] Exhibit 1 separates this column into decisions logged as standard "revise-and-resubmit" outcomes and decisions logged as "uncertain" outcomes that allow but do not necessarily encourage resubmission.

[d] Total acceptances and conditional acceptances include manuscripts published and forthcoming in *The Accounting Review*. Thus, the total acceptances during a year do not represent the actual number of articles published during that year. Acceptance totals include only those files accepted from the normal review process, and exclude invited commentaries from AAA Presidential Scholar Lectures and from occasional research forums.

to-issue basis, meaning that the articles accepted for publication at the time of the deadline for the next issue became the articles published in that issue. This process can create imbalances due to the sometimes cyclical nature of the flow of submissions and revisions, so I decided early on that I would like to have one or two issues "in waiting" at the time of each issue deadline, providing a buffer to help smooth out the cyclical variation. Accordingly, during the last half of 2008, my coeditors and I built a buffer of approximately 15–20 forthcoming articles at the conditional acceptance stage, and we have held that buffer roughly constant ever since.

A second reason for the downturn in acceptances from 81 last year to 63 this year is a bit embarrassing, but I will provide full disclosure. In May 2010, I encountered a personal backlog, due to my end-of-semester teaching responsibilities, among other priorities. Thus, some conditional acceptance letters drafted by my coeditors in April or May did not get sent out until June, as I always read and provide my own comments (and usually a marked copy) on all accepted manuscripts before sending out the decision. Ever the good accountant, I strictly applied a May 31 cutoff in preparing Table 2, but the reader should be aware that we sent out ten additional acceptance letters in the first half of June 2010. Including those ten would have brought the fiscal 2010 acceptance total to 73 instead of 63. The good news is that I get to count them next year.

To put the number of acceptances in perspective, the reader should consider our capacity constraint. As long as journals continue to publish bound hard-copies, there are only so many articles that can fit in a bound issue. Effective 2008, *TAR* has published six times per year. In my experience, I can fit up to 12 articles in each issue. Thus, if we fill the journal to capacity, we have room for $12 \times 6 = 72$ articles per year, one of which goes automatically to the Presidential Scholar Lecture. For the most part, we have been filling the journal to its practical capacity, which happily coincides with what I perceive to be a reasonable, productive acceptance rate that publishes good research while maintaining a high-quality standard. I cannot make any promises about our ability to sustain full-capacity production, as an editor never knows what is coming next, but I think the journal is in good health.

Table 2, Panel B: Final Outcome Resolution for All New Submissions

New this year is supplemental Panel B to Table 2, reflecting what I think is an excellent suggestion from AAA Publications Committee member Bob Kaplan to calculate the journal's "acceptance rate" in a different way. As I understand the suggestion, if one views any given year's new submissions as the population from which articles can be accepted, then in subsequent years, after most of the revisions have been processed, one can tally how many of that year's new submissions were accepted, how many were rejected, and how many remain pending due to outstanding invitations to revise. Table 2, Panel B reports this supplemental format for the 557 new submissions received during the journal year ending May 31, 2009.[2] Of these, 45 (8.1 percent) have been accepted through May 31, 2010, 442 (79.4 percent) have been rejected, and the remaining 70 manuscripts (12.6 percent) are still outstanding as of May 31, 2010 due to pending revisions. These statistics provide a useful floor and ceiling for *TAR*'s acceptance rate: for fiscal 2008–2009, the acceptance rate can be no lower than 8.1 percent, as those manuscripts have already been accepted, and can be no higher than 20.6 percent if all remaining pending revisions from fiscal 2008–2009 are eventually accepted. Of course, it is unrealistic to assume that either none or all of the pending revisions will be accepted. My best estimate is that splitting the difference between the two numbers (an estimated acceptance rate of 14.4 percent) comes mate-

[2] Table 2, Panel B does not report outcomes for new submissions during the journal year ending May 31, 2010, as too many of those new submissions are awaiting a decision as of May 31, 2010 to make the statistics meaningful.

rially close to our true acceptance rate for the first year of operations under the current editorial regime. I will provide a follow-up table next year.

Exhibit 1: Histogram of Editorial Rounds and Outcomes

Exhibit 1 provides a histogram of decision outcomes by round for the journal year ending May 31, 2010. As I did last year, I am taking the liberty of reporting more detail than that requested by the Committee, showing separate bars for "revise" and "uncertain" invitations as well as separate bars for "contribution" and "validity" rejections. The 475 first-round decisions in Panel A of Exhibit 1 differ from the 502 new submissions reported in Panel A of Table 1 because Exhibit 1 includes 98 new submissions "in process" as of June 1, 2009 and excludes 125 new submissions for which a decision had not yet been reached as of May 31, 2010.

Exhibit 1, Panel A shows that most first-round manuscripts are rejected because they lack a sufficient incremental contribution for *The Accounting Review*. Ever since we started logging rejection decisions as being based primarily on contribution issues or validity issues, I have been struck by the fact that the former category is more than twice as frequent as the latter. Indeed, my sense of the modal first-round review report that recommends rejection is that it contains wording similar to, "While this study appears to have been competently executed, it does not provide much new insight relative to what we already know from the extensive prior literature in this area." Put simply, following the bandwagon in a well-researched area is not necessarily a path to success in a top-tier journal.

Under the current editorial regime, *The Accounting Review* has employed two revision categories: (1) standard "revise-and-resubmit" letters, and (2) "uncertain" letters. Both outcomes allow the author to revise, but an "uncertain" letter indicates that the outcome risk is higher than usual for an invitation to revise, such that the path to a successful revision is unclear. When sending an "uncertain" letter, we encourage a careful assessment of whether the author sees an effective way to address the concerns raised, and if not, that submission elsewhere might be best for the author from a cost-benefit perspective. The easiest way for an author to tell which kind of letter s/he has received is that an "uncertain" letter asks for an email reply indicating whether the author intends to revise and resubmit. While some might reason that an author should always exercise an option to revise, the honest intent of an "uncertain" letter is that the decision is unclear from the editor's perspective, as a multiple-round rejection can be worse for the author than a first-round rejection. As Exhibit 1 indicates, most revision letters are of the standard "revise-and-resubmit" variety, but we try to use the more cautious "uncertain" wording when appropriate. To avoid misunderstanding, I should clarify that even a standard "revise-and-resubmit" letter conveys outcome risk; it is only the degree of that risk that differs between the two revision categories.

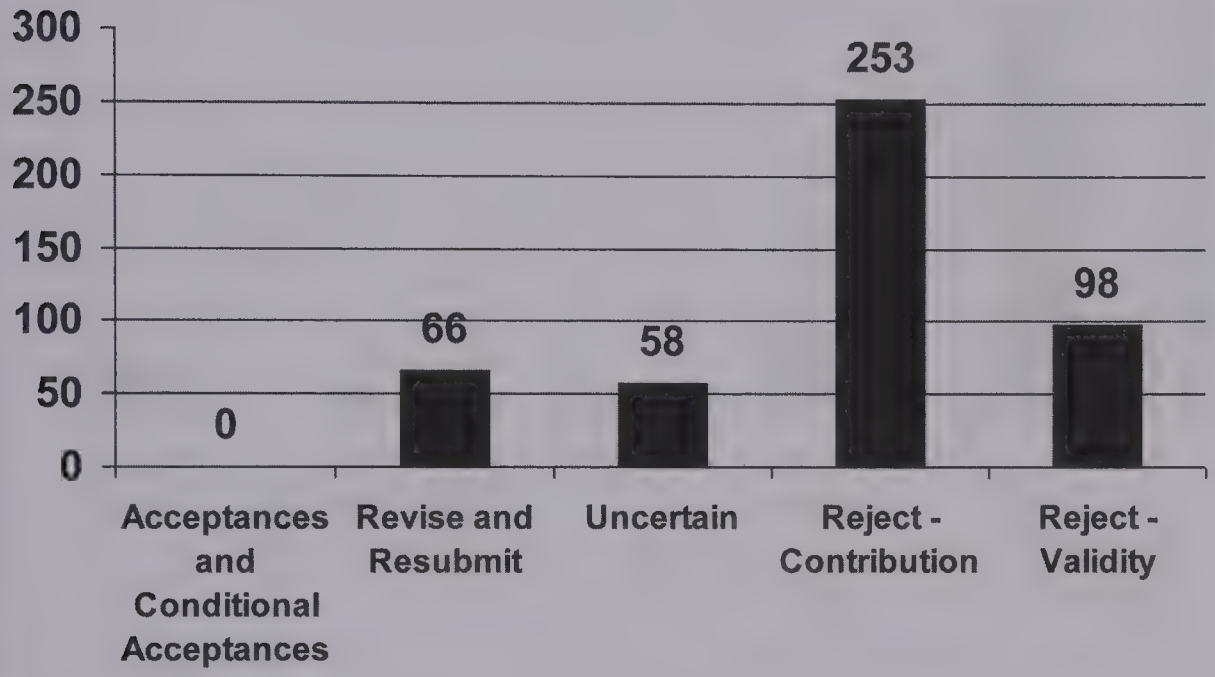
Exhibit 1 indicates that the relative odds of success increase substantially in the second round and beyond. By the time a manuscript reaches the third or fourth round (Panel C), most outcomes are acceptances, as should be the case for manuscripts that have advanced to this stage. Still, nothing is guaranteed, as is evidenced by the seven rejections in Exhibit 1, Panel C. As an aside, the vast majority of the decisions tallied in Exhibit 1, Panel C are for third-round manuscripts, with only nine that went to the fourth round. We did have one fifth-round manuscript within the journal year ending May 31, 2010, which, fortunately, was accepted for publication.

Table 3: Submissions and Acceptances by Subject Area and Research Method

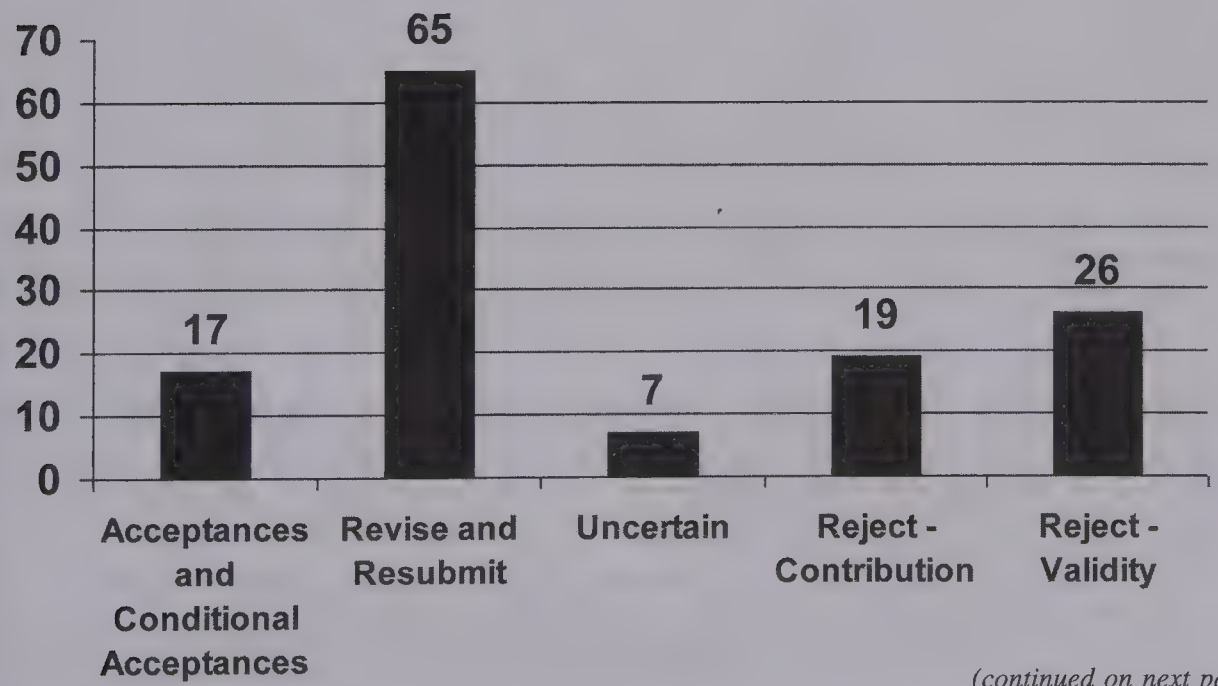
Perhaps the most interesting statistics are in Table 3, which categorizes submissions and acceptances by primary subject area (Panel A), primary method (Panel B), and subject areas crossed with methods (Panel C). Very similar to last year, all three panels show that the proportion of acceptances by area is nearly identical to the proportion of submissions by area. Contingency table tests fail to reject the null hypothesis of equivalent submission and acceptance proportions by

EXHIBIT 1
Histogram of Editorial Rounds and Outcomes
Journal Year Ending May 31, 2010

Panel A: First-Round Outcomes (New Submissions)



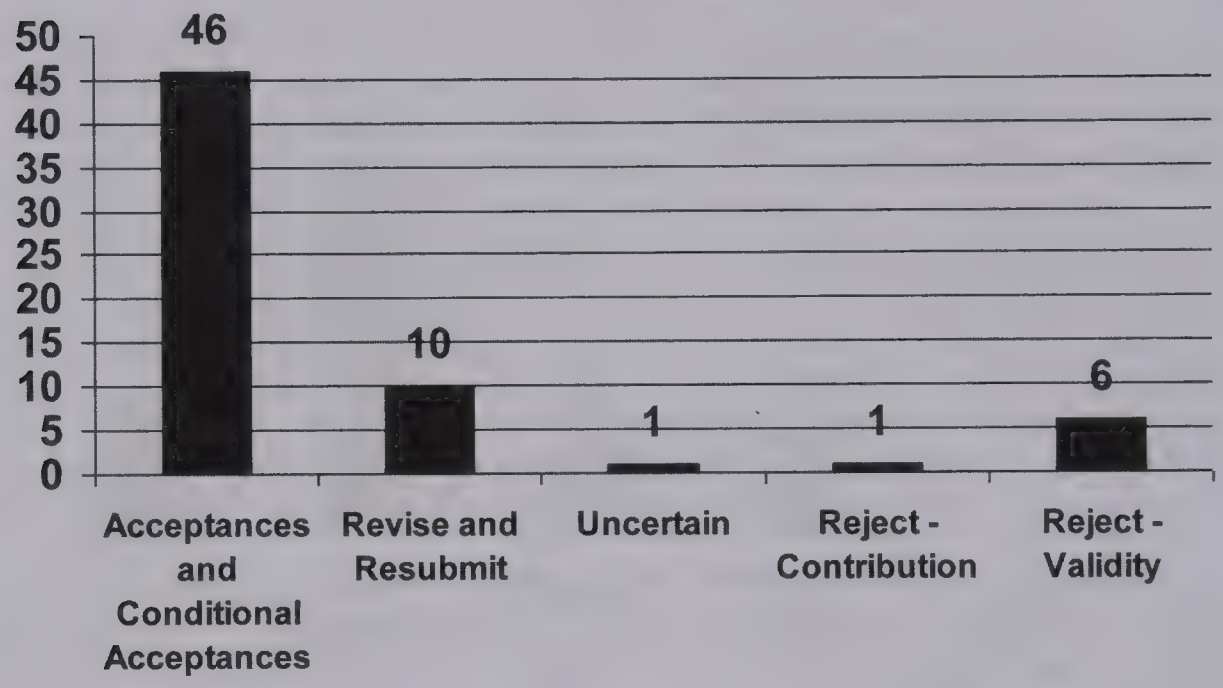
Panel B: Second-Round Outcomes (First Revisions)



(continued on next page)

EXHIBIT 1

Panel C: Third- and Fourth-Round Outcomes (Second and Third Revisions)



subject area ($\chi^2_{df=7} = 5.74$; $p > 0.50$) or by method ($\chi^2_{df=5} = 6.35$; $p = 0.27$).[3] As was the case last year, if any area might be able to state a case for underrepresentation relative to submissions, it would be those who conduct financial-archival research (45 percent of submissions and 37 percent of acceptances). That said, a journal editor would probably take issue with that statement (justifiably) if this editorial were subjected to the review process, given the lack of statistical significance as reported above.

The broader point is that one cannot draw inferences about acceptance rates by area (a ratio) by looking at what journals publish (the numerator of that ratio).[4] Table 3 sheds insight by reporting our denominators in terms of submissions by area, and I submit that these data add an important qualification to Tuttle and Dillard's (2007, 398) assertion that the research and publication process in the United States is "privileging financial accounting research competing for

[3] To address the possibility that small cell sizes in the less frequent categories might be distorting the statistics, I repeated both the subject area and methodological contingency table analyses after combining the "governmental/NFP," "international," "systems," and "other" subject areas into a combined "other" subject area category and after combining the "field/case," "survey," and "other" methods into a combined "other" methodological category. Results continue to fail to reject the null hypothesis of equivalent submission and acceptance percentages for both subject areas ($\chi^2_{df=4} = 4.09$; $p = 0.39$), and methods ($\chi^2_{df=3} = 4.09$; $p = 0.25$).

[4] Merchant (2010) is a classic case in point, drawing inferences about various journal specializations (and implied biases) from the nature of the articles they publish. One cannot deny that different accounting journals publish different frequencies of articles in various areas, as Merchant (2010) demonstrates persuasively, but analyses of journal content cannot differentiate the possibility of editorial biases from the broader phenomenon of self-selection by submitting authors.

TABLE 3

Submissions and Acceptances by Subject Area and Research Method
Journal Year Ending May 31, 2010

Panel A: Submissions and Acceptances by Subject Area

Primary Subject Area (a)	Submissions Representing Unique Manuscript Files with Decisions (b)	Percentage of Total Submissions (c)	Acceptances and Conditional Acceptances (d)	Percentage of Total Acceptances (e)
Auditing	110	17.9%	13	20.6%
Financial	327	53.2%	31	49.2%
Governmental and Not-for-Profit	7	1.1%	0	0.0%
International	37	6.0%	2	3.2%
Managerial	84	13.7%	8	12.7%
Systems	12	2.0%	2	3.2%
Taxation	34	5.5%	7	11.1%
Other	4	0.6%	0	0.0%
Total	615	100.0%	63	100.0%

Panel B: Submissions and Acceptances by Method

Primary Research Method (a)	Submissions Representing Unique Manuscript Files with Decisions (b)	Percentage of Total Submissions (c)	Acceptances and Conditional Acceptances (d)	Percentage of Total Acceptances (e)
Analytical	42	6.8%	7	11.1%
Empirical-Archival	457	74.3%	40	63.5%
Experimental	86	14.0%	13	20.6%
Field and Case Study	8	1.3%	2	3.2%
Survey	15	2.4%	1	1.6%
Other	7	1.1%	0	0.0%
Total	615	100.0%	63	100.0%

(continued on next page)

Panel C: Submissions and (Acceptances) by Subject Area Crossed with Method

Subject and Method	Auditing	Financial	Managerial	Taxation	All Other Subjects	Total
Analytical	7 (1)	20 (3)	12 (2)	1 (1)	2 (0)	42 (7)
	1% (2%)	3% (5%)	2% (3%)	<1% (2%)	<1% (0%)	7% (11%)
Empirical-Archival	70 (5)	276 (23)	36 (3)	29 (6)	46 (3)	457 (40)
	11% (8%)	45% (37%)	6% (5%)	5% (10%)	7% (5%)	74% (63%)
Experimental	28 (6)	25 (5)	22 (2)	2 (0)	9 (0)	86 (13)
	5% (10%)	4% (8%)	4% (3%)	<1% (0%)	1% (0%)	14% (21%)
All Other Methods	5 (1)	6 (0)	14 (1)	2 (0)	3 (1)	30 (3)
	1% (2%)	1% (0%)	2% (2%)	<1% (0%)	<1% (2%)	5% (5%)
Total	110 (13)	327 (31)	84 (8)	34 (7)	60 (4)	615 (63)
	18% (21%)	53% (49%)	14% (13%)	6% (11%)	10% (6%)	100%

In Panels A and B:

(a) Subject (Panel A) and method (Panel B) categories are as directed by the AAA Publications Committee. Manuscripts that overlapped categories were assigned to the category best capturing the study's primary objective, as judged by the Senior Editor.

(b) This column reflects the 615 unique manuscript files obtained by taking the 673 editorial decisions during the fiscal year as tallied in Table 1, Panel A and subtracting the 58 revise-and-resubmit decisions on manuscripts for which a subsequent decision was logged within the same fiscal year on the same file, thereby avoiding double counting of categories for the same manuscript files.

(c) This column represents the entry in Column (b) divided by 615 unique manuscript files.

(d) This column reflects the 63 acceptance and conditional acceptance decisions reached on the 615 unique manuscript files in Column (b).

(e) This column represents the entry in Column (d) divided by 63 acceptances and conditional acceptances.

In Panel C:

Cell entries reflect submissions first, then acceptances (in parentheses). The top row for each cell indicates raw counts. The bottom row computes percentages of the 615 total unique submissions and 63 total acceptances, respectively, rounded to the nearest whole percentage to enable the table to fit in the available space.

prestige journal space." While Tuttle and Dillard (2007) (and others) may well be correct in asserting that financial-archival research is commanding an ever larger share of the pie, Table 3 suggests that *The Accounting Review* is a reflection of that phenomenon, not the cause.

In raising this point over the past year regarding the similar statistics reported in my previous annual report (Kachelmeier 2009), I have often heard the counterargument that submissions by area are endogenous (to use an empirical-archival term), reflecting authors' preferences and beliefs regarding their likely prospects at any given journal. Put simply, researchers in underrepresented areas will not submit to *The Accounting Review* if they perceive that they will not get a fair shake at *The Accounting Review*. I understand this argument, but I also find it to be circular. It is difficult for me to understand how any journal can be criticized for not publishing research it does not receive.

To be sure, inclusive journals should take steps to signal openness to the scholarly community, such as through its editorial board appointments and editors. I will leave the reader to evaluate the diversity of our 129 editorial board members and the additional 453 *ad hoc* reviewers named in the Appendix. As for editors, I am joined by a diverse set of outstanding coeditors such as Shannon Anderson (expertise in field studies), Jim Hunton (expertise in accounting information systems), and Tom Omer (expertise in taxation). The rest seems up to the community. If, at some point, a future Table 3 shows an imbalance between submissions by area and acceptances by area, then we can engage a different debate as to the causes of that imbalance. But until then, I think Table 3 indicates that we are running a fair game, with all areas having roughly equal prospects of success. As the acceptance-rate data indicate, those prospects are not very high *ex ante*. But the same high standards of quality and rigor apply to all.

Table 4: Author Affiliations

Table 4, Panel A tallies the number of articles corresponding to each university for which at least one (co)author published an article in Volume 84 (calendar 2009) or Volume 85 (calendar 2010) of *The Accounting Review*. Articles written by k coauthors are attributed 1/k to each author's affiliation, and if an author lists two affiliations, half of that author's 1/k share is allocated to each affiliation. As I did last year, I list the institutions in Table 4, Panel A in descending order of the number of articles published.[5] However, the intent of Panel A is most certainly not to compare universities in terms of research productivity. A recent study by Coyne et al. (2010) reports a more comprehensive analysis along that line. Rather, the point of Table 4, Panel A is to document the diversity of *TAR*'s authors. Fully 130 different institutions were represented in 136 research articles published in *The Accounting Review* over the past two years. That seems pretty diverse to me.

Table 4, Panel B reports data on international diversity, tallying frequency counts for both submissions and acceptances for U.S. authors and non-U.S. authors. Separate columns report these data for the journal year ending May 31, 2010 and for the two-year cumulative totals combining this year and last, with the aggregate statistics reducing the cumulative submission count for files submitted in both years due to revision. Authors from non-U.S. affiliations comprise 27.0 percent (26.7 percent) of the fiscal 2010 (cumulative) submissions, and 20.1 percent (19.8 percent) of the acceptances. A contingency table test does not reject the (two-tailed) null hypothesis of equal submission and acceptance percentages of U.S. versus non-U.S. authors for fiscal 2010 ($\chi^2_{df=1} = 1.19$; $p = 0.28$), but the statistic becomes marginally significant in the more powerful cumula-

[5] The fact that The University of Texas at Austin happens to be listed first this year (6.08 articles, or 4.5 percent of the two-year total) gives me a good opportunity to clarify that, by policy, a *TAR* coeditor never serves as editor on one of his/her own colleague's manuscripts. The 6.08 coauthor-adjusted *TAR* articles for The University of Texas at Austin over the past two years reflect nine different authors across a variety of topics and methods.

TABLE 4

Author Affiliation

Panel A: Universities and Other Institutions Represented in Vols. 84 (2009) and 85 (2010)

Author Affiliation	Current-Year Number of Published Articles (Vol. 85, 2010)	Two-Year Cumulative Number of Published Articles (2009 and 2010)
The University of Texas at Austin	5.58	6.08
University of Pennsylvania	2.25	4.67
Stanford University	2.17	4.17
Michigan State University	1.75	3.92
Harvard University	1.58	3.75
Dartmouth College	3.42	3.42
University of Illinois at Urbana–Champaign	1.50	3.17
The University of Georgia	2.92	2.92
The University of Iowa	0.50	2.75
University of Michigan	1.33	2.67
University of Pittsburgh	1.00	2.67
University of Toronto	0.83	2.58
Indiana University	1.00	2.50
The University of Chicago	1.50	2.33
The George Washington University	1.00	2.33
Florida International University	1.25	2.08
The Pennsylvania State University	1.33	2.00
The University of New South Wales	0.00	1.83
University of Missouri–Columbia	1.00	1.83
Nanyang Technological University	1.50	1.83
Hong Kong Polytechnic University	1.67	1.79
The University of Oklahoma	1.17	1.75
The University of Texas at Dallas	0.58	1.75
University of California, Los Angeles	1.33	1.67
Emory University	1.67	1.67
The University of North Carolina at Chapel Hill	1.33	1.67
Texas A&M University	0.33	1.67
The University of Utah	0.83	1.67
University of Wisconsin–Madison	0.25	1.58
University of California, Berkeley	0.00	1.50
Singapore Management University	0.00	1.50
University of Minnesota	0.00	1.25
University of Alberta	1.17	1.17
Brigham Young University	0.00	1.17
Hong Kong University of Science and Technology	0.33	1.17
University of Southern California	0.00	1.17
University of Virginia	0.50	1.17
Aarhus University	1.00	1.00
University of Arkansas	1.00	1.00
Universidad Carlos III de Madrid	0.00	1.00
Georgia Institute of Technology	0.00	1.00
Humboldt University–Berlin	0.00	1.00

(continued on next page)

Panel A: Universities and Other Institutions Represented in Vols. 84 (2009) and 85 (2010)

Author Affiliation	Current-Year Number of Published Articles (Vol. 85, 2010)	Two-Year Cumulative Number of Published Articles (2009 and 2010)
Iowa State University	0.00	1.00
McGill University	0.00	1.00
University of Melbourne	0.00	1.00
Miami University (Ohio)	1.00	1.00
National University of Singapore	1.00	1.00
University of New Hampshire	1.00	1.00
Purdue University	1.00	1.00
University of Rochester	0.00	1.00
Saint Louis University	1.00	1.00
Santa Clara University	1.00	1.00
Southern Methodist University	1.00	1.00
Tel Aviv University	1.00	1.00
Washington University at St. Louis	0.00	1.00
University of Alabama	0.83	0.83
Cornell University	0.33	0.83
University of Houston	0.33	0.83
Maastricht University	0.00	0.83
Massachusetts Institute of Technology	0.67	0.83
The Ohio State University	0.83	0.83
University of Oregon	0.83	0.83
Seoul National University	0.58	0.83
Bentley University	0.25	0.75
University of California, Irvine	0.00	0.75
Erasmus University	0.33	0.75
Duke University	0.00	0.67
Rice University	0.00	0.67
University of South Florida	0.00	0.67
University of Washington	0.67	0.67
Concordia University	0.50	0.63
University of British Columbia	0.33	0.58
University of Connecticut	0.00	0.58
University of Notre Dame	0.00	0.58
University of South Carolina	0.58	0.58
University of Amsterdam	0.00	0.50
Boston University	0.50	0.50
Columbia University	0.00	0.50
University of Florida	0.00	0.50
Goethe University	0.00	0.50
Korea University	0.50	0.50
University of Miami	0.50	0.50
Norwegian School of Management	0.50	0.50
Utah State University	0.00	0.50
Washington State University	0.00	0.50
University of Waterloo	0.00	0.50

(continued on next page)

Panel A: Universities and Other Institutions Represented in Vols. 84 (2009) and 85 (2010)

Author Affiliation	Current-Year Number of Published Articles (Vol. 85, 2010)	Two-Year Cumulative Number of Published Articles (2009 and 2010)
University of Antwerp	0.00	0.33
University of Arizona	0.33	0.33
Arizona State University	0.00	0.33
University of Auckland	0.00	0.33
University of California, Davis	0.00	0.33
University of California, San Diego	0.00	0.33
Carnegie Mellon University	0.33	0.33
China Europe International Business School	0.33	0.33
Colorado State University	0.00	0.33
Drexel University	0.33	0.33
George Mason University	0.33	0.33
Gradient Analytics, Inc.	0.33	0.33
HEC Paris	0.00	0.33
Inha University	0.33	0.33
Lehigh University	0.33	0.33
University of Massachusetts Dartmouth	0.33	0.33
McMaster University	0.33	0.33
Meijo University	0.00	0.33
University of Nevada, Las Vegas	0.00	0.33
New York University	0.00	0.33
North Carolina State University	0.33	0.33
Northeastern University	0.00	0.33
Northwestern University	0.33	0.33
Osaka University of Economics	0.00	0.33
Sungkyunkwan University	0.33	0.33
Syracuse University	0.00	0.33
The University of Tennessee	0.00	0.33
Virginia Commonwealth University	0.33	0.33
Virginia Polytechnic Institute and State University	0.33	0.33
Wichita State University	0.33	0.33
Ball State University	0.00	0.25
Chapman University	0.00	0.25
University of Hong Kong	0.00	0.25
Illinois State University	0.25	0.25
Indian School of Business	0.00	0.25
Kent State University	0.00	0.25
University of Massachusetts, Amherst	0.25	0.25
Mississippi State University	0.25	0.25
Oklahoma State University	0.25	0.25
Tilburg University	0.25	0.25
Barclays Global Investors	0.00	0.17
Copenhagen Business School	0.00	0.17
Interdisciplinary Center–Herzliya	0.00	0.17
University of Southern Maine	0.00	0.17

(continued on next page)

Panel B: U.S. and non-U.S. Affiliations Represented in Vols. 84 (2009) and 85 (2010)

	Journal Year Ending May 31, 2010		Cumulative from June 1, 2008 to May 31, 2010	
	Number (Percentage) of Submissions Representing Unique Manuscript Files with Decisions	Number (Percentage) of Acceptances	Number (Percentage) of Submissions Representing Unique Manuscript Files with Decisions	Number (Percentage) of Acceptances
U.S. Institutions	448.93 (73.0%)	50.34 (79.9%)	825.84 (73.3%)	115.50 (80.2%)
Non-U.S. Institutions	166.07 (27.0%)	12.66 (20.1%)	301.16 (26.7%)	28.50 (19.8%)

Panel A reports author affiliations for all articles published in Volume 84 (calendar 2009) and Volume 85 (calendar 2010) of *The Accounting Review*, excluding invited commentaries. For articles written by k coauthors, the table allocates 1/k of an article to each coauthor's affiliation. For authors indicating two affiliations, that author's 1/k share is allocated equally to both affiliations.

Panel B reports data for authors' affiliations, allocating 1/k of each manuscript to each of k coauthors. Submissions data for the journal year ending May 31, 2010 reflect the 615 unique files for which decisions were reached in fiscal 2010, excluding files for which two or more decisions were reached within the year due to invitations to revise and resubmit. Cumulative submissions data from June 1, 2008 to May 31, 2010 (i.e., two journal years) reflect the 1,127 unique manuscript files for which a decision was logged in either year, but not double counting files for which a decision was reached in both years (due to revision). Acceptance data reflect all acceptances and conditional acceptances reached in the journal years ending May 31, 2009 and May 31, 2010, including some acceptances that are scheduled for publication in early 2011. Hence, the acceptance data in this panel will not equal the total acceptances in Panel A, as Panel A only reflects actual articles published in calendar 2009 and calendar 2010, excluding forthcoming articles accepted for future publication.

tive test ($\chi^2_{df=1} = 3.01$; $p = 0.08$). My hunch is that non-U.S. authors are probably underrepresented a bit in acceptances relative to submissions, but as I indicated last year, I honestly do not perceive any such underrepresentation as an overt bias against non-U.S. authors. Rather, I suspect that the factors that influence reviewer assessments of manuscript quality are correlated to some extent with the U.S./non-U.S. categorization. To avoid misunderstanding, please note that this conjecture is a comment on the distribution, not a generalization about non-U.S. authors. For the vast majority of our submissions, I have discerned no systematic quality differences between U.S. and non-U.S. authors, and indeed, current data indicate that about one out of every five articles published in *TAR* is written by a non-U.S. author. My comment applies only to the tails of the distributions.

Table 5: Processing Time

Table 5 reports processing time data for the 673 decisions reached from June 1, 2009 through May 31, 2010, defining "processing time" as the number of days from the submission to the date my assistant, Mary Capps, sends the decision letter. Our overall processing time has not improved since last year—in fact, the average is up from 81 days last year to 86 days this year. I was curious to learn why it takes nearly three months to reach an (average) editorial decision, so I dug a bit deeper by first determining the *later* reviewer for each reviewed submission and then calculating the average review time for that later reviewer. Calculating the review time for the later of the two reviewers is appropriate because a two-reviewer system is subject to somewhat of a "weakest-link" problem—even if the first reviewer is timely, a decision based on two reviews cannot proceed until the second review arrives.

For the journal year ending May 31, 2010, the later of the two reviewers took an average of 55 days to return his/her review—call it two months. If our average total turnaround time is about three months, this means that the typical manuscript is out for review for two months, with an

TABLE 5

Processing Time from Date of Submission to Date of Sending the Decision
Journal Year Ending May 31, 2010

Processing Time	Number of Manuscripts	Percentage	Cumulative Number of Manuscripts	Cumulative Percentage
≤30 days	36	5.3%	36	5.3%
31–60 days	72	10.7%	108	16.0%
61–90 days	227	33.7%	335	49.8%
91–120 days	303	45.0%	638	94.8%
≥121 days	35	5.2%	673	100.0%

Mean processing time: 86 days.
Median processing time: 91 days.

additional month for (1) initial administrative processing to log the submission in our system and check for previous, related submissions, (2) the initial "pre-review" by my research assistant and by me to identify potential reviewers, (3) reviewer clearance with the assigned editor, (4) processing invitations to both reviewers and following up, as necessary, (5) finding and sending invitations to alternative reviewer(s) if the initially invited reviewer(s) decline, (6) editorial processing of the reviews, (7) editorial processing of the manuscript, (8) formulating and writing the decision letter, (9) senior editor reading of a coeditor's decision letter and consultation, if applicable, and (10) administrative processing of the decision letter and corresponding statistics for our databases. Yes, I know this ten-point list sounds a bit defensive, but my intent is to communicate that there are many things to do for each file, and especially when other demands (like teaching) are present, it is difficult to get the total non-review time materially lower than a month, though we try and are successful to varying degrees at different times of the year. (The Holiday season is especially problematic—authors probably should not expect a very quick turnaround when submitting a manuscript on December 15.)

Regarding review time, some might advise reaching a decision based on one review if the second review is not in hand within two months, and indeed, that was my intent when I started the term. Especially this year, however, my coeditors and I have tended to wait for both reviews, even if that means a turnaround time that exceeds what we like to see. It is easy to favor a one-reviewer decision in abstract terms, but when one is looking at an actual manuscript with one negative review that could be due to a matter of taste, I think most authors would favor waiting a couple more weeks for a second opinion. *Ex post*, of course, if the later review also turns out negative and the decision is to reject, it would have been better to have made that decision earlier, but editors do not have the luxury of hindsight when only one review is in hand.

There are a couple of bright spots in our turnaround statistics. First, while many manuscripts take three months from submission to decision, very few (5.2 percent to be exact) take four months or longer. Also, if the goal is to reach a final decision on a timely basis, I would submit that the turnaround time for each round comprises only a small part of the total. Time spent on the author's desk between rounds can be quite significant, as can the total number of rounds before a final decision is reached. Along that line, I think it is comforting that only nine of our 673 editorial decisions in fiscal 2010 were in the fourth round, and only one reached the fifth round (upon which it was accepted). We try to reach closure by the third round in the vast majority of cases.

IV. SOME NOTES OF THANKS AND REMEMBRANCE

Notes of Thanks

As I did last year, I close this report with several notes of thanks. First, I thank the person with whom most submitting authors and reviewers are quite familiar but have never met—Mary Capps. Mary works tirelessly to keep the "trains running on time," and she also is the source of those emails many reviewers have received with the capitalized word "REMINDER" at the beginning of the subject line. She is also fond of the one-line email "Have you had a chance to get to this yet?," followed by a smiley face icon. At a recent conference I attended, a faculty member observed to me, "Mary is quite persistent, isn't she?" He said it with a smile, and quickly followed up with a note of appreciation that someone cares so much about making the journal work as efficiently and effectively as possible, while remaining sensitive to personal circumstances that can arise from time to time. I agree.

Second, I thank my doctoral research assistant, Tracie Majors. Tracie has become "scary good" at researching new submissions to identify potential reviewers. She recognizes potential conflicts of interest not only from affiliations and coauthorships, but also from different perspectives. A typical Tracie comment would be, "Professor X clearly knows this area, but the author is taking issue with X's favored conclusion, so if you decide to go with X, you really need a second reviewer with a more independent perspective." Some might question the appropriateness of a doctoral student playing such an active role in proposing possible reviewers, but I view it as one of the wisest decisions I have made as senior editor. Unlike professors who develop different kinds of political baggage, Tracie's only exposure to a new submission is the manuscript itself and the literature to which it contributes. To be clear, Tracie only suggests a list of potential reviewers for my consideration (she does not make the decisions), but I have found her input to be invaluable.

Third, I thank my coeditors, Shannon Anderson, Harry Evans, Jim Hunton, Kathryn Kadous, Sanjay Kallapur, Ranjani Krishnan, Laureen Maines, Paul Newman, Tom Omer, Wayne Thomas, Mark Trombley, Beverly Walther, and Paul Zarowin. They never complain, even when they should, and their decisions are consistently of the highest ethical and professional caliber. For all their hard work, they get paid the grand sum of zero, so I am extremely grateful for their dedicated service. With over 500 new submissions per year, it would be impossible for any senior editor to do this job alone. Delegation to experts is essential, and I am comfortable placing complete trust in these 13 outstanding professionals.

Fourth, I thank *TAR*'s Editorial Advisory and Review Board, as listed in the inside cover material of each issue. Like our coeditors, Editorial Board members get paid the whopping sum of zero for their service to *TAR*, and they put in a tremendous amount of hours completing six and sometimes more reviews per year. Six reviews per year might not seem too oppressive, but one has to keep in mind that our most demanded reviewers are probably also in high demand by other journals, and indeed, several of them serve on multiple journal editorial boards. In essence, the reward for gaining a reputation for doing timely, high-quality reviews is more review requests. We let no good deed go unpunished.

Fifth, I thank our *ad hoc* reviewers, the unsung heroes of any successful journal. Continuing a tradition I started last year, the Appendix lists all 453 *ad hoc* reviewers who completed one or more review reports between June 1, 2009 and May 31, 2010. I challenge the reader to find any research area (topical or methodological) that is not covered by at least a few names in this Appendix, so I hope that, in addition to thanking these individuals for their generosity, the Appendix also serves to signal *TAR*'s diversity.

Sixth, I thank the exceptional professional staff of the American Accounting Association. I would start naming names, but surely I would inadvertently leave someone out who should be thanked, so I will just leave it at that. We send them accepted manuscripts, and they magically

make a neat looking journal appear, including an invaluable copyediting service. (Authors: please do not read that as an excuse not to check your conditionally accepted manuscripts carefully before sending us the final revision.)

Seventh, I thank The University of Texas at Austin, including McCombs School of Business Dean Tom Gilligan and Department of Accounting Chair Urton Anderson, for supporting my service as senior editor of *The Accounting Review*. While most of my working hours since 2008 have been for an activity governed by the American Accounting Association, it is The University of Texas at Austin that pays my salary and makes this service possible. I still teach, which is my preference; I would never want to be out of the classroom entirely. But it is clear that the nature of my academic position has changed dramatically over the past two years (actually three, counting the extensive planning efforts an incoming senior editor must undertake before the term begins). I thank my employer for understanding and supporting me in this role and for recognizing the requisite tradeoffs.

Eighth, and perhaps most importantly, I thank my wife Paula and my daughters Karen, Nicole, and Sara. Last year at this time, my annual report mentioned the deck lighting I had been stalling for months while working virtually every weekend on *TAR* matters. I finally got the lighting installed (I think I could faintly hear my daughters in the background singing the Hallelujah chorus). This year's winner of the procrastination award goes to my 1992 Camaro, which sits in our garage gathering spider webs while I stall taking action on either fixing it or putting it out of its misery. We are all going camping next week—this time in Colorado. I will put a few *TAR* files in the pile with the camping gear, which I hope Paula "forgets" to bring along (meaning the files, not the gear). So add seven days to our average turnaround statistics for this month—I need the break, and family comes first when all is said and done. One sad note is that our loyal dog Midnight will not be anxiously awaiting our return this time, as she passed away last spring at the ripe old age of 11. The nice thing about dogs is that their love is unconditional, no matter how badly you screw up. We miss her greatly.

Remembrances

While not part of the annual report per se, I will exercise my editorial prerogative to offer a few comments on two leaders in our profession who passed away this year—John Dickhaut and Anthony Hopwood. Both exhibited inspirational leadership in daring to challenge the received wisdom in accounting research, and both were true scholars in every sense of that word.

My most vivid early recollection of John Dickhaut was when I was a doctoral student at the University of Florida in the mid-1980s and John was an invited workshop guest. I distinctly remember an intense discussion of backward induction and multiperiod reputations with John while in the men's underwear department of the local JCPenney, as John had asked me to take him there during the lunch hour because he had forgotten to pack any clothes for his two-day visit. This was vintage John Dickhaut—a man who could forget trivial things like packing a suitcase, but who at the same time was absolutely brilliant in his understanding of the interface between human behavior and accounting. About ten years later, I gave John Dickhaut most of the credit for making sure the two words "and experimental" were included in the "Aims and Scope" statement for *Review of Accounting Studies*, a new journal for which John was one of the founding editors. For *The Accounting Review*, I take special pride in the fact that we had the privilege of publishing two of John's final scholarly works—both in the same issue (Dickhaut 2009; Dickhaut and Xin 2009). Those who value diversity and quality in accounting scholarship should read both articles. Finally, on a somewhat sad but telling note, the reader will see John's name in the Appendix of *ad hoc* reviewers this year, as he turned in a review of a first-round submission just the week before he passed away, using voice-recognition software to help compose his thoughts. He recommended revision.

I did not have as much personal knowledge of Anthony Hopwood, but I greatly appreciate his influence on the discipline, not only as editor of *Accounting, Organizations and Society*, but also as a persistent champion for diversity in accounting scholarship. I wonder how he might react to Table 3 in this report. In the commentary based on his 2006 AAA Presidential Scholar Lecture, Professor Hopwood issued a special challenge for *The Accounting Review* as the Association's flagship journal, urging that "every effort be made to encourage *The Accounting Review* to embrace the new, the innovative, what accounting research might be in the process of becoming, and new interdisciplinary perspectives … in a way that provides both a catalyst and a model for other journals of influence" (Hopwood 2007, 1373). The challenge is profound. I am trying, and it is my heartfelt conviction that *TAR* senior editors and coeditors past, present, and future would join me in agreeing with Professor Hopwood's challenge. But innovation in any discipline is a team sport. We need the help of the scholarly community in accounting to make it happen.

APPENDIX

TAR Ad Hoc Reviewers, June 1, 2009–May 31, 2010

Note: This table lists all *ad hoc* reviewers who submitted at least one review report between June 1, 2009 and May 31, 2010. It does not include the 129 members of the Editorial Advisory and Review Board, as those names are listed separately in the inside cover material of each issue. *Ad hoc* reviewers whose reports were submitted after May 31, 2010 will be included in next year's report.

Lawrence Abbott, University of Wisconsin–Milwaukee
A. Rashad Abdel-khalik, University of Illinois at Urbana–Champaign
Margaret Abernethy, University of Melbourne
David Aboody, University of California, Los Angeles
Andrew Acito, The University of Iowa
Christopher Agoglia, University of Massachusetts Amherst
Kris Allee, Michigan State University
Michael Alles, Rutgers University
Urton Anderson, The University of Texas at Austin
Christopher Armstrong, University of Pennsylvania
Sharad Asthana, The University of Texas at San Antonio
Brad Badertscher, University of Notre Dame
Kee-Hong Bae, York University
Mark Bagnoli, Purdue University
Bok Baik, Seoul National University
Wendy Bailey, University of South Carolina
Steven Balsam, Temple University
Michael Bamber, The University of Georgia
Ran Barniv, Kent State University
Mary Barth, Stanford University
Jan Barton, Emory University
Eli Bartov, New York University
Mark Beasley, North Carolina State University
Paul Beck, University of Illinois at Urbana–Champaign
Joy Begley, University of British Columbia
Bruce Behn, University of Tennessee
Messod Beneish, Indiana University
Daniel Bens, University of Arizona

Jeremy Bertomeu, Northwestern University
Anne Beyer, Stanford University
Sreedhar Bharath, Arizona State University
Neil Bhattacharya, Southern Methodist University
Sanjeev Bhojraj, Cornell University
Gary Biddle, The University of Hong Kong
Bruce Billings, Florida State University
Mary Brooke Billings, New York University
Erv Black, Brigham Young University
Jeff Boone, The University of Texas at San Antonio
Jan Bouwens, Tilburg University
Robert Bowen, University of Washington
Kendall Bowlin, University of Mississippi
Joe Brazel, North Carolina State University
Francois Brochet, Harvard University
Jason Brown, Indiana University
Jennifer Brown, Arizona State University
Jörg Budde, University of Bonn
David Burgstahler, University of Washington
Jeffrey Burks, University of Notre Dame
Robert Bushman, University of North Carolina at Chapel Hill
Brian Cadman, University of Utah
Steven Cahan, University of Auckland
Andrew Call, The University of Georgia
Dennis Campbell, Harvard University
Eddy Cardinaels, Tilburg University
Peter Carey, Monash University
Tina Carpenter, The University of Georgia
Elizabeth Carson, The University of New South Wales
Judson Caskey, University of California, Los Angeles
Gavin Cassar, University of Pennsylvania
Jeffrey Casterella, Colorado State University
Marcus Caylor, University of South Carolina
Sandra Chamberlain, University of British Columbia
Dennis Chambers, Kennesaw State University
Chih-Ying Chen, Singapore Management University
Clara Xiaoling Chen, University of Illinois at Urbana–Champaign
Kung Chen, University of Nebraska—Lincoln
Peter Chen, Hong Kong University of Science and Technology
Qi Chen, Duke University
Shuping Chen, The University of Texas at Austin
Xia Chen, University of Wisconsin—Madison
Agnes Cheng, Louisiana State University
Shijun Cheng, University of Maryland
Jong-Hag Choi, Seoul National University
Margaret Christ, The University of Georgia
Peter Christensen, Aarhus University
Ted Christensen, Brigham Young University
Wai-Fong Chua, The University of New South Wales

Hyeesoo (Sally) Chung, Arizona State University West
Bryan Church, Georgia Institute of Technology
Peter Clarkson, University of Queensland
Shana Clor-Proell, University of Wisconsin—Madison
Daniel Cohen, New York University
Dan Collins, The University of Iowa
Carlos Corona, The University of Texas at Austin
Masako Darrough, Baruch College–CUNY
Somnath Das, University of Illinois at Chicago
Sudipto Dasgupta, Hong Kong University of Science and Technology
Angela Davis, University of Oregon
Gus De Franco, University of Toronto
Henri Dekker, Vrije University Amsterdam
Joel Demski, University of Florida
Aiyesha Dey, University of Minnesota
Ilia Dichev, Emory University
John Dickhaut, Chapman University
J. Richard Dietrich, The Ohio State University
Shane Dikolli, Duke University
Ming Dong, York University
Jeffrey Doyle, Utah State University
Julia D'Souza, Cornell University
Scott Dyreng, Duke University
Christine Earley, Providence College
David Easley, Cornell University
Frank Ecker, Duke University
Merle Ederhof, University of Michigan
Eti Einhorn, Tel Aviv University
Randal Elder, Syracuse University
W. Brooke Elliott, University of Illinois at Urbana–Champaign
David Erkens, University of Southern California
Michael Ettredge, University of Kansas
Diana Falsetta, University of Miami
Qintao Fan, University of California, Berkeley
C. Edward Fee, Michigan State University
Bill Felix, University of Arizona
Mei Feng, University of Pittsburgh
Andrew Ferguson, University of Technology Sydney
Fabrizio Ferri, New York University
Rebecca Files, The University of Texas at Dallas
Urs Fischbacher, University of Konstanz
Joseph Fisher, Indiana University
Mary Margaret Frank, University of Virginia
Richard Frankel, Washington University at St. Louis
Pingyang Gao, The University of Chicago
Jon Garfinkel, The University of Iowa
Jennifer Gaver, The University of Georgia
Lisa Milici Gaynor, University of South Florida
Weili Ge, University of Washington

Marshall Geiger, University of Richmond
Aloke Ghosh, Baruch College–CUNY
Michael Gibbs, The University of Chicago
Dan Givoly, The Pennsylvania State University
Guojin Gong, The Pennsylvania State University
Elizabeth Gordon, Temple University
Severin Grabski, Michigan State University
Jeffery Gramlich, University of Southern Maine
Audrey Gramling, Kennesaw State University
Barbara Murray Grein, Drexel University
Paul Griffin, University of California, Davis
Zhaoyang Gu, University of Minnesota
Wayne Guay, University of Pennsylvania
Ferdinand Gul, The Hong Kong Polytechnic University
Sanjay Gupta, Michigan State University
Ilan Guttman, Stanford University
Charles Hadlock, Michigan State University
Jane Hamilton, LaTrobe University
Jackie Hammersley, The University of Georgia
Michelle Hanlon, Massachusetts Institute of Technology
Bruce Haslem, Florida State University
John Hassell, Indiana University Indianapolis
Richard Hatfield, University of Alabama
Rachel Hayes, University of Utah
Carla Hayn, University of California, Los Angeles
Gary Hecht, Emory University
Frank Heflin, Florida State University
Bill Heninger, Brigham Young University
Karen Hennes, University of Oklahoma
Don Herrmann, Oklahoma State University
Max Hewitt, Indiana University
Stephen Hillegeist, INSEAD
Jessen Hobson, University of Illinois at Urbana–Champaign
Leslie Hodder, Indiana University
Vicky Hoffman, University of Pittsburgh
Rani Hoitash, Bentley University
Udi Hoitash, Northeastern University
Lori Holder-Webb, Western New England College
Keith Houghton, Australian National University
Paul Hribar, The University of Iowa
John Hughes, University of California, Los Angeles
Kai Wai Hui, Hong Kong University of Science and Technology
Mingyi Hung, University of Southern California
David Hurtt, Baylor University
Mark Huson, University of Alberta
Amy Hutton, Boston College
Lee-Seok Hwang, Seoul National University
Yuhchang Hwang, Arizona State University
Troy Hyatt, Boise State University

Paul Irvine, The University of Georgia
Alan Jagolinzer, Stanford University
Karim Jamal, University of Alberta
Surya Janakiriman, The University of Texas at Dallas
Nicole Jenkins, Vanderbilt University
Ross Jennings, The University of Texas at Austin
Kevan Jensen, University of Oklahoma
John (Xuefeng) Jiang, Michigan State University
W. Bruce Johnson, The University of Iowa
Derek Johnston, Colorado State University
Bjorn Jorgensen, University of Colorado at Boulder
Paul Kalyta, McGill University
Sok-Hyon Kang, The George Washington University
Tony Kang, Oklahoma State University
Ron Kasznik, Stanford University
Sharon Katz, Columbia University
Elizabeth Keating, Boston College
Mozaffar Khan, University of Minnesota
Jeong-Bon Kim, City University of Hong Kong
Oliver Kim, University of Maryland
Michael Kinney, Texas A&M University
Sandy Klasa, University of Arizona
Ken Klassen, University of Waterloo
Kalin Kolev, Yale University
Gopal Krishnan, Lehigh University
Jagan Krishnan, Temple University
Linda Krull, University of Oregon
Xi (Jason) Kuang, Georgia Institute of Technology
Susan Kulp, The George Washington University
Soo Young Kwon, Korea University
Ryan LaFond, BlackRock, Inc.
Richard Lambert, University of Pennsylvania
Wayne Landsman, University of North Carolina at Chapel Hill
Mark Lang, University of North Carolina at Chapel Hill
Christian Laux, Goethe University
Volker Laux, The University of Texas at Austin
Charles Lee, Stanford University
Yen-Jung Lee, National Taiwan University
Craig Lefanowicz, University of Virginia
Clive Lennox, Nanyang Technological University
Andrew Leone, University of Miami
Christian Leuz, The University of Chicago
Baruch Lev, New York University
Edward X. Li, University of Rochester
Laura Yue Li, University of Illinois at Urbana–Champaign
Oliver Li, University of Arizona
Siqi Li, Santa Clara University
Xu Li, Lehigh University
Yue Li, University of Toronto

Pierre Liang, Carnegie Mellon University
Woody Liao, University of California, Riverside
Anne Lillis, University of Melbourne
Marlys Lipe, University of Oklahoma
Robert Lipe, University of Oklahoma
Petro Lisowsky, University of Illinois at Urbana–Champaign
Chao-Shin Liu, University of Notre Dame
Jing Liu, University of California, Los Angeles
Josh Livnat, New York University
Kin Lo, University of British Columbia
Gerald Lobo, University of Houston
Tom Lopez, University of Alabama
Ken Lorek, Northern Arizona University
Tim Loughran, University of Notre Dame
Henock Louis, The Pennsylvania State University
Tong Lu, University of Houston
Yvonne Lu, Lehigh University
Russell Lundholm, University of Michigan
Luann Lynch, University of Virginia
Victor Maas, University of Amsterdam
Mario Maletta, Northeastern University
Xiumin Martin, Washington University at St. Louis
Michal Matějka, Arizona State University
Zoltan Matolcsy, University of Technology Sydney
Ella Mae Matsumura, University of Wisconsin–Madison
Steven Matsunaga, University of Oregon
Elaine Mauldin, University of Missouri–Columbia
Bill Mayew, Duke University
Cheri Mazza, Fordham University
Mary Lea McAnally, Texas A&M University
Annie McGowan, Texas A&M University
Sean McGuire, Texas A&M University
John McInnis, The University of Texas at Austin
James McKeown, The Pennsylvania State University
Maureen McNichols, Stanford University
Nahum Melumad, Columbia University
Krishnopal Menon, Boston University
Molly Mercer, DePaul University
Richard Mergenthaler, The University of Iowa
William Messier, University of Nevada, Las Vegas
Brian Miller, Indiana University
Jeffrey Miller, University of Notre Dame
Birendra Mishra, University of California, Riverside
Stephen Moehrle, University of Missouri–St. Louis
Frank Moers, Maastricht University
Emad Mohammad, McMaster University
Kimberly Moreno, Northeastern University
Richard Morton, Florida State University
Karl Muller, The Pennsylvania State University

Linda Myers, University of Arkansas
Mark Myring, Ball State University
Nandu Nagarajan, University of Pittsburgh
Vic Naiker, The University of Auckland
Dhananjay Nanda, University of Miami
Gordian Ndubizu, Drexel University
Alexander Nekrasov, University of California, Irvine
Mark Nelson, Cornell University
Andrew Newman, University of Pittsburgh
Jeffrey Ng, Massachusetts Institute of Technology
Terence Ng, Nanyang Technological University
D. Craig Nichols, Cornell University
Christine Nolder, Bentley University
David North, University of Richmond
Patricia O'Brien, University of Waterloo
Maria Ogneva, Stanford University
Steven Orpurt, Pepperdine University
Jaime Ortega, Universidad Carlos III de Madrid
Linda Parsons, University of Alabama
Evelyn Patterson, Indiana University Indianapolis
Kenneth Peasnell, Lancaster University
Raynolde Pereira, University of Missouri–Columbia
Gary Peters, University of Arkansas
Michael Peters, Villanova University
Christine Petrovits, New York University
Ray Pfeiffer, Texas Christian University
Marc Picconi, Indiana University
Arianana Pinello, Georgia State University
Matt Pinnuck, University of Melbourne
Mini Pizzini, Southern Methodist University
George Plesko, University of Connecticut
Elizabeth Plummer, Texas Christian University
Susan Porter, University of Virginia
Gordon Potter, Cornell University
Grace Pownall, Emory University
Doug Prawitt, Brigham Young University
Madhav Rajan, Stanford University
Dasaratha Rama, Florida International University
Kartik Raman, Bentley University
K. Ramesh, Rice University
Sunderesh Ramnath, University of Miami
Srinivasan Rangan, University of California, Davis
Bill Rankin, Colorado State University
Susan Ravenscroft, Iowa State University
Korok Ray, Georgetown University
David Reeb, Temple University
Lynn Rees, Texas A&M University
Sonja Olhoft Rego, The University of Iowa
Ken Reichelt, Louisiana State University

Robert Resutek, Dartmouth College
J. Kenneth Reynolds, Indiana University
David Ricchiute, University of Notre Dame
Jay Rich, Illinois State University
Ed Riedl, Harvard University
Andrea Alston Roberts, University of Virginia
Leslie Robinson, Dartmouth College
Jonathan Rogers, The University of Chicago
Jacob Rose, University of New Hampshire
Joshua Rosett, Claremont McKenna College
Brian Rountree, Rice University
Casey Rowe, Purdue University
Sugata Roychowdhury, Boston College
Tjomme Rusticus, Northwestern University
Stephen Ryan, New York University
Gil Sadka, Columbia University
Steve Salterio, Queen's University
Maria Sanchez, Rider University
Tatiana Sandino, University of Southern California
Srinivasan Sankaraguruswamy, National University of Singapore
Richard Saouma, University of California, Los Angeles
Haresh Sapra, The University of Chicago
Andrew Schmidt, Columbia University
Susan Scholz, University of Kansas
Joseph Schultz, Arizona State University
Steven Schwartz, Binghamton University, SUNY
Dan Segal, Interdisciplinary Center Herzliya
Frank Selto, University of Colorado at Boulder
Partha Sengupta, George Mason University
Nick Seybert, University of Maryland
Doug Shackelford, University of North Carolina at Chapel Hill
Ron Shalev, Washington University at St. Louis
Philip Shane, University of Colorado at Boulder
Divesh Sharma, Florida International University
Kenneth Shaw, University of Missouri–Columbia
Min Shen, George Mason University
Terry Shevlin, University of Washington
Michael Shields, Michigan State University
Jae Yong Shin, Seoul National University
Pervin Shroff, University of Minnesota
Stephanie Sikes, University of Pennsylvania
Paul Simko, University of Virginia
Roger Simnett, The University of New South Wales
Dan Simunic, University of British Columbia
Jason Smith, University of Nevada, Las Vegas
Michael Smith, Boston University
Naomi Soderstrom, University of Colorado at Boulder
Brian Spilker, Brigham Young University
Sri Sridhar, Northwestern University

Dhinu Srinivasan, University of Pittsburgh
Suraj Srinivasan, Harvard University
Anup Srivastava, Northwestern University
Paul Steinbart, Arizona State University
Doug Stevens, Florida State University
Don Stokes, Monash University
Roland Strausz, Humboldt University–Berlin
Stephen Stubben, University of North Carolina at Chapel Hill
Scott Summers, Brigham Young University
Amy Sun, The Pennsylvania State University
Yan Sun, Saint Louis University
Jayanthi Sunder, Northwestern University
Shyam V. Sunder, Northwestern University
Shyam Sunder, Yale University
Ed Swanson, Texas A&M University
Vicki Tang, Georgetown University
Bill Tayler, Emory University
Siew-Hong Teoh, University of California, Irvine
Jane Thayer, The University of Georgia
Robert Trezevant, University of Southern California
Jennifer Wu Tucker, University of Florida
Wim Van der Stede, London School of Economics
Laurence van Lent, Tilburg University
Ann Vanstraelen, Maastricht University
Florin Vasvari, London Business School
Uma Velury, University of Delaware
Ramgopal Venkataraman, Southern Methodist University
Raghu Venugopalan, University of Illinois at Urbana–Champaign
Rodrigo Verdi, Massachusetts Institute of Technology
Robert Verrecchia, University of Pennsylvania
Cynthia Vines, University of Kentucky
Gnanakumar Visvanathan, George Mason University
Alfred Wagenhofer, University of Graz
James Wahlen, Indiana University
Dechun Wang, Texas A&M University
Isabel Wang, Michigan State University
Shiing-Wu Wang, University of Southern California
Xue (Sue) Wang, Emory University
Gregory Waymire, Emory University
Connie Weaver, Texas A&M University
Alan Webb, University of Waterloo
Dan Weiss, Tel Aviv University
Michael Welker, Queen's University
Scott Whisenant, University of Houston
Sally Widener, Rice University
Heather Wier, University of Alberta
T. Jeffrey Wilks, Brigham Young University
Michael Willenborg, University of Connecticut
David Williams, The Ohio State University

Michael Williams, Gradient Analytics, Inc.
Michael Williamson, The University of Texas at Austin
Ryan Wilson, The University of Iowa
Wendy Wilson, Southern Methodist University
Jennifer Winchel, University of South Carolina
Christopher Wolfe, Texas A&M University
Franco Wong, University of Toronto
Donghui Wu, Hong Kong Polytechnic University
Joanna Wu, University of Rochester
Martin Wu, University of Illinois at Urbana–Champaign
Anne Wyatt, University of Queensland
Jingyoung Park Wynn, Louisiana Tech University
Peter Wysocki, University of Miami
Hong Xie, University of Kentucky
David Yermack, New York University
Michelle Yetman, University of California, Davis
Robert Yetman, University of California, Davis
Ping Eric Yeung, The University of Georgia
Han Yi, University of Oklahoma
Lance Young, University of Washington
Yong Yu, The University of Texas at Austin
Tzachi Zach, The Ohio State University
Guochang Zhang, Hong Kong University of Science and Technology
Ivy Zhang, University of Minnesota
Jieying Zhang, University of Southern California
May H. Zhang, University of Missouri–Columbia
Ping Zhang, University of Toronto
X. Frank Zhang, Yale University
Yuan Zhang, Columbia University
Yue May Zhang, Northeastern University
Yun Zhang, Duke University
Yijiang Zhao, American University
Amir Ziv, Columbia University

REFERENCES

Blank, R. M. 1991. The effects of double-blind versus single-blind reviewing: Experimental evidence from *The American Economic Review*. *The American Economic Review* 81 (December): 1041–1067.

Coyne, J. G., S. L. Summers, B. Williams, and D. A. Wood. 2010. Accounting program research rankings by topical area and methodology. *Issues in Accounting Education* 25 (4).

Dickhaut, J. 2009. The brain as the original accounting institution. *The Accounting Review* 84 (November): 1703–1712.

——, and B. Xin. 2009. Market efficiencies and drift: A computational model. *The Accounting Review* 84 (November): 1805–1831.

Gilliland, S. W., and J. M. Cortina. 1997. Reviewer and editor decision making in the journal review process. *Personnel Psychology* 50 (Summer): 427–452.

Hopwood, A. G. 2007. Whither accounting research? *The Accounting Review* 82 (October): 1365–1374.

Kachelmeier, S. J. 2004. Reviewing the review process. *The Journal of the American Taxation Association* 26 (Supplement): 143–154.

———. 2009. Annual report and editorial commentary for *The Accounting Review*. *The Accounting Review* 84 (November): 2047–2075.

Lynch, J. G., Jr. 1998. Reviewing. *Advances in Consumer Research* 25: 1–6.

Merchant, K. A. 2010. Paradigms in accounting research: A view from North America. *Management Accounting Research* 21 (June): 116–120.

Tuttle, B., and J. Dillard. 2007. Beyond competition: Institutional isomorphism in U.S. accounting research. *Accounting Horizons* 21 (December): 387–409.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature. Topical areas of interest to the journal include accounting information systems, auditing and assurance services, financial accounting, management accounting, taxation, and all other areas of accounting, broadly defined. The journal is also open to all rigorous research methods.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and editors are responsible for providing constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary.*

FORMAT

1. All manuscripts should be formatted in 12-point font on 8 1/2 x 11″ paper and should be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins should be at least one inch from top, bottom, and sides.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we".
5. A cover page should show the title of the paper, all authors' names, titles and affiliations, email addresses, and any acknowledgments.
6. The American Accounting Association encourages use of gender-neutral language in its publications.

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a cross-sectional equation; re-form*. See *Webster's* for correct usage.

Keywords: The abstract must be followed by at least three keywords to assist in indexing the paper and identifying qualified reviewers.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words (150 maximum) should be presented on a separate page immediately preceeding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. The Keywords statement should appear immediately below the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. When the reference list contains two or more works by the same author (as the only author or first of two or more authors) published in the same year, the suffix a, b, etc., is appended to the date in the within-text citations and in the "References" section. For example, (Johansson 2004a, 2004b, 2004c; Baiman and Rajan 2002a, 2002b; Dhaliwal et al. 2005a, 2005b).
4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works should use acronyms or short titles where practicable. For example: (NCFFR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in *alphabetical order* according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Use authors' initials instead of proper names.
3. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
4. Date of publication follows the author's (authors') name(s).
5. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

———, and ———. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.

Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: http://hsgac.senate.gov/files/111803berry.pdf.

Cohen, D., A. Dey, and T. Lys. 2005. The Sarbanes Oxley Act of 2002: Implications for compensation structure and risk-taking incentives of CEOs. Working paper, New York University, University of Chicago, and Northwestern University.

Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition.* Ann Arbor, MI: University of Michigan.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

———, and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.

Dhaliwal, D., M. Erickson, and O. Li. 2005a. Shareholder income taxes and the relation between earnings and returns. *Contemporary Accounting Research* 22: 587–616.

———, L. Krull, O. Li, and W. Moser. 2005b. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43: 675–708.

Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.

Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109.* FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.

Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes.* Boston, MA: Houghton Mifflin.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.

Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics* (Winter) 30: 555–574.

National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: NCFFR.

Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.

———. 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy*, edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.

Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.

Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports.* Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be inserted using the "footnote" or "endnote" feature of the word processing software which will automatically number the footnotes throughout the manuscript with superscript Arabic numerals.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state upon submission that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. Manuscripts should be submitted via email as Microsoft Word or PDF file to the Senior Editor, Steven Kachelmeier, at email address tar@mccombs.utexas.edu. Please submit separate files for (1) the manuscript's title page with identifying information (not forwarded to reviewers), (2) the manuscript with title page and all other identifying information removed, and (3) any necessary supplemental files, such as experimental instructions and/or response memoranda on invited revisions. The body or subject line of the email should also indicate the submission fee receipt code from AAA (see step 4 below).
4. A nonrefundable submission fee is required unless explicitly waived by the editor for invited revisions of previous submissions. The submission fee in U.S. funds is $200.00 for members and $400.00 for nonmembers of the AAA payable by credit card (VISA or MasterCard only). The payment form is available online at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If you are unable to pay by credit card or have any questions, please contact the AAA Member Services Team at (941) 921-7747 or info@aaahq.org.
5. Revisions must be submitted within 12 months from the decision letter inviting a revision.

COMMENTS

The Accounting Review welcomes submissions of comments on previous *TAR* articles. Comments on articles previously published in *The Accounting Review* will generally be reviewed by two reviewers, usually including an author of the original article (to assist the editor in evaluating whether the submitted comment represents the prior article accurately) and an additional reviewer who is independent of the original article. If a comment is accepted for publication, the original author will generally be invited to reply. All other editorial requirements, as enumerated above, apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The AAA's Executive Committee policy (originally adopted in 1989, and amended in 2009) is that the objective of the Association-wide journals (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. To fulfill this objective, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles.

AUTHOR INDEX TO VOLUME 85

Author	*Title*	*Page*
AHN, TAE SIK (and HWANG and KIM)	The Impact of Performance Measure Discriminability on Ratee Incentives	389
ARMSTRONG, CHRISTOPHER S. (and BARTH and JAGOLINZER and RIEDL)	Market Reaction to the Adoption of IFRS in Europe	31
BAGNOLI, MARK (and WATTS)	Oligopoly, Disclosure, and Earnings Management	1191
BAIMAN, STANLEY (and NETESSINE and SAOUMA)	Informativeness, Incentive Compensation, and the Choice of Inventory Buffer	1839
BAMBER, E. MICHAEL (and HAMMERSLEY and CARPENTER)	The Influence of Documentation Specificity and Priming on Auditors' Fraud Risk Assessments and Evidence Evaluation Decisions	547
BAMBER, LINDA SMITH (and HUI and YEUNG)	Managers' EPS Forecasts: Nickeling and Diming the Market?	63
BAMBER, LINDA SMITH (and JIANG and PETRONI and WANG)	Comprehensive Income: Who's Afraid of Performance Reporting?	97
BAMBER, LINDA SMITH (and JIANG and WANG)	What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure	1131
BARTH, MARY E. (and ARMSTRONG and JAGOLINZER and RIEDL)	Market Reaction to the Adoption of IFRS in Europe	31
BARTON, JAN (and HANSEN and POWNALL)	Which Performance Measures Do Investors Around the World Value the Most—and Why?	753
BARUA, ABHIJIT (and FAN and CREADY and THOMAS)	Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items	1303
BARUA, ABHIJIT (and LIN and SBARAGLIA)	Earnings Management Using Discontinued Operations	1485
BEATTY, ANNE (and LIAO and WEBER)	Financial Reporting Quality, Private Information, Monitoring, and the Lease-versus-Buy Decision	1215
BLOUIN, JENNIFER L. (and GLEASON and MILLS and SIKES)	Pre-Empting Disclosure? Firms' Decisions Prior to FIN No. 48	791

BOL, JASMIJN C. (and KEUNE and MATSUMURA and SHIN)	Supervisor Discretion in Target Setting: An Empirical Investigation	1861
BOWEN, ROBERT M. (and CALL and RAJGOPAL)	Whistle-Blowing: Target Firm Characteristics and Economic Consequences	1239
BRAZEL, JOSEPH F. (and CARPENTER and JENKINS)	Auditors' Use of Brainstorming in the Consideration of Fraud: Reports from the Field	1273
BROCHET, FRANCOIS	Information Content of Insider Trades before and after the Sarbanes-Oxley Act	419
CADMAN, BRIAN (and KLASA and MATSUNAGA)	Determinants of CEO Pay: A Comparison of ExecuComp and Non-ExecuComp Firms	1511
CALL, ANDREW C. (and BOWEN and RAJGOPAL)	Whistle-Blowing: Target Firm Characteristics and Economic Consequences	1239
CARPENTER, TINA D. (and HAMMERSLEY and BAMBER)	The Influence of Documentation Specificity and Priming on Auditors' Fraud Risk Assessments and Evidence Evaluation Decisions	547
CARPENTER, TINA D. (and BRAZEL and JENKINS)	Auditors' Use of Brainstorming in the Consideration of Fraud: Reports from the Field	1273
CASKEY, JUDSON (and NAGAR and PETACCHI)	Reporting Bias with an Audit Committee	447
CASSAR, GAVIN (and GERAKOS)	Determinants of Hedge Fund Internal Controls and Fees	1887
CHEN, SHIMIN (and SUN and WU)	Client Importance, Institutional Improvements, and Audit Quality in China: An Office and Individual Auditor Level Analysis	127
CHRISTENSEN, JOHN	Accounting Errors and Errors of Accounting	1827
CHRISTENSEN, PETER O. (and DE LA ROSA and FELTHAM)	Information and the Cost of Capital: An *Ex Ante* Perspective	817
CHRISTENSEN, PETER O. (and ŞABAC and TIAN)	Ranking Performance Measures in Multi-Task Agencies	1545
CREADY, WILLIAM M. (and FAN and BARUA and THOMAS)	Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items	1303
CREADY, WILLIAM (and LOPEZ and SISNEROS)	The Persistence and Market Valuation of Recurring Nonrecurring Items	1577
D'SOUZA, JULIA M. (and RAMESH and SHEN)	The Interdependence between Institutional Ownership and Information Dissemination by Data Aggregators	159

DE LA ROSA, LEONIDAS E. (and CHRISTENSEN and FELTHAM)	Information and the Cost of Capital: An *Ex Ante* Perspective	817
DYRENG, SCOTT D. (and HANLON and MAYDEW)	The Effects of Executives on Corporate Tax Avoidance	1163
EDERHOF, MERLE	Discretion in Bonus Plans	1921
FAN, YUN (and BARUA and CREADY and THOMAS)	Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items	1303
FELTHAM, GERALD A. (and CHRISTENSEN and DE LA ROSA)	Information and the Cost of Capital: An *Ex Ante* Perspective	817
FENG, MEI (and KOCH)	Once Bitten, Twice Shy: The Relation between Outcomes of Earnings Guidance and Management Guidance Strategy	1951
FENG, MEI (and McVAY)	Analysts' Incentives to Overweight Management Guidance When Revising Their Short-Term Earnings Forecasts	1617
FISCHER, PAUL E. (and STOCKEN)	Analyst Information Acquisition and Communication	1985
GAO, PINGYANG	Disclosure Quality, Cost of Capital, and Investor Welfare (2008 American Accounting Association Competitive Manuscript Award Winner)	1
GERAKOS, JOSEPH (and CASSAR)	Determinants of Hedge Fund Internal Controls	1887
GIVOLY, DAN (and HAYN and KATZ)	Does Public Ownership of Equity Improve Earnings Quality?	195
GLEASON, CRISTI A. (and BLOUIN and MILLS and SIKES)	Pre-Empting Disclosure? Firms' Decisions Prior to FIN No. 48	791
GOLD, ANNA (and HUNTON)	A Field Experiment Comparing the Outcomes of Three Fraud Brainstorming Procedures: Nominal Group, Round Robin, and Open Discussion	911
GOODWIN, JOHN (and GUL)	Short-Term Debt Maturity Structures, Credit Ratings, and the Pricing of Audit Services	877
GOW, IAN D. (and ORMAZABAL and TAYLOR)	Correcting for Cross-Sectional and Time-Series Dependence in Accounting Research	483
GUENTHER, DAVID A. (and SANSING)	The Effect of Tax-Exempt Investors and Risk on Stock Ownership and Expected Returns	849

GUL, FERDINAND A. (and GOODWIN)	Short-Term Debt Maturity Structures, Credit Ratings, and the Pricing of Audit Services	877
GUTTMAN, ILAN	The Timing of Analysts' Earnings Forecasts	513
HAMMERSLEY, JACQUELINE S. (and BAMBER and CARPENTER)	The Influence of Documentation Specificity and Priming on Auditors' Fraud Risk Assessments and Evidence Evaluation Decisions	547
HANLON, MICHELLE (and DYRENG and MAYDEW)	The Effects of Executives on Corporate Tax Avoidance	1163
HANSEN, THOMAS BOWE (and BARTON and POWNALL)	Which Performance Measures Do Investors Around the World Value the Most—and Why?	753
HATFIELD, RICHARD C. (and HOUSTON and STEFANIAK and USREY)	The Effect of Magnitude of Audit Difference and Prior Client Concessions on Negotiations of Proposed Adjustments	1647
HAYN, CARLA K. (and GIVOLY and KATZ)	Does Public Ownership of Equity Improve Earnings Quality?	195
HOPE, OLE-KRISTIAN (and LANGLI)	Auditor Independence in a Private Firm and Low Litigation Risk Setting	573
HOUSTON, RICHARD W. (and HATFIELD and STEFANIAK and USREY)	The Effect of Magnitude of Audit Difference and Prior Client Concessions on Negotiations of Proposed Adjustments	1647
HUGHES, JOHN S. (and WILLIAMS and LEVINE)	Influence of Capital Gains Tax Policy on Credibility of Unverified Disclosures	719
HUI, KAI WAI (and BAMBER and YEUNG)	Managers' EPS Forecasts: Nickeling and Diming the Market?	63
HUNTON, JAMES E. (and GOLD)	A Field Experiment Comparing the Outcomes of Three Fraud Brainstorming Procedures: Nominal Group, Round Robin, and Open Discussion	911
HWANG, INY (and AHN and KIM)	The Impact of Performance Measure Discriminability on Ratee Incentives	389
JAGOLINZER, ALAN D. (and ARMSTRONG and BARTH and RIEDL)	Market Reaction to the Adoption of IFRS in Europe	31
JAMAL, KARIM (and TAN)	Joint Effects of Principles-Based versus Rules-Based Standards and Auditor Type in Constraining Financial Managers' Aggressive Reporting	1325
JENKINS, J. GREGORY (and BRAZEL and CARPENTER)	Auditors' Use of Brainstorming in the Consideration of Fraud: Reports from the Field	1273

JIANG, JOHN (XUEFENG) (and BAMBER and PETRONI and WANG)	Comprehensive Income: Who's Afraid of Performance Reporting? 97
JIANG, JOHN (XUEFENG) (and BAMBER and WANG)	What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure 1131
KACHELMEIER, STEVEN J. (and WILLIAMSON)	Attracting Creativity: The Initial and Aggregate Effects of Contract Selection on Creativity-Weighted Productivity 1669
KANAGARETNAM, KIRIDARAN (and KRISHNAN and LOBO)	An Empirical Analysis of Auditor Independence in the Banking Industry 2011
KANG SOK-HYON (and ZHAO)	Information Content and Value Relevance of Depreciation: A Cross-Industry Analysis 227
KATZ, SHARON P. (and GIVOLY and HAYN)	Does Public Ownership of Equity Improve Earnings Quality? 195
KEUNE, TIMOTHY M. (and BOL and MATSUMURA and SHIN)	Supervisor Discretion in Target Setting: An Empirical Investigation 1861
KEUNG, EDMUND C.	Do Supplementary Sales Forecasts Increase the Credibility of Financial Analysts' Earnings Forecasts? 2047
KIM, DONGCHEOL (and QI)	Accruals Quality, Stock Returns, and Macroeconomic Conditions 937
KIM, MYUNG-IN (and AHN and HWANG)	The Impact of Performance Measure Discriminability on Ratee Incentives 389
KLASA, SANDY (and CADMAN and MATSUNAGA)	Determinants of CEO Pay: A Comparison of ExecuComp and Non-ExecuComp Firms 1511
KOCH, ADAM S. (and FENG)	Once Bitten, Twice Shy: The Relation between Outcomes of Earnings Guidance and Management Guidance Strategy 1951
KOONCE, LISA (and WILLIAMSON and WINCHEL)	Consensus Information and Nonprofessional Investors' Reaction to the Revelation of Estimate Inaccuracies 979
KRISHNAN, GOPAL V. (and KANAGARETNAM and LOBO)	An Empirical Analysis of Auditor Independence in the Banking Industry 2011
LANGLI, JOHN CHRISTIAN (and HOPE)	Auditor Independence in a Private Firm and Low Litigation Risk Setting 573
LAUX, VOLKER (and NEWMAN)	Auditor Liability and Client Acceptance Decisions 261

LEONE, ANDREW J. (and LIU)	Accounting Irregularities and Executive Turnover in Founder-Managed Firms	287
LEVINE, CAROLYN B. (and WILLIAMS and HUGHES)	Influence of Capital Gains Tax Policy on Credibility of Unverified Disclosures	719
LI, SIQI	Does Mandatory Adoption of International Financial Reporting Standards in the European Union Reduce the Cost of Equity Capital?	607
LIAO, SCOTT (and BEATTY and WEBER)	Financial Reporting Quality, Private Information, Monitoring, and the Lease-versus-Buy Decision	1215
LIN, STEVE (and BARUA and SBARAGLIA)	Earnings Management Using Discontinued Operations	1485
LISOWSKY, PETRO	Seeking Shelter: Empirically Modeling Tax Shelters Using Financial Statement Information	1693
LIU, MICHELLE (and LEONE)	Accounting Irregularities and Executive Turnover in Founder-Managed Firms	287
LOBO, GERALD J. (and KANAGARETNAM and KRISHNAN)	An Empirical Analysis of Auditor Independence in the Banking Industry	2011
LOPEZ, THOMAS J. (and CREADY and SISNEROS)	The Persistence and Market Valuation of Recurring Nonrecurring Items	1577
MASLI, ADI (and PETERS and RICHARDSON and SANCHEZ)	Examining the Potential Benefits of Internal Control Monitoring Technology	1001
MATSUMURA, ELLA MAE (and BOL and KEUNE and SHIN)	Supervisor Discretion in Target Setting: An Empirical Investigation	1861
MATSUNAGA, STEVE (and CADMAN and KLASA)	Determinants of CEO Pay: A Comparison of ExecuComp and Non-ExecuComp Firms	1511
MAYDEW, EDWARD L. (and DYRENG and HANLON)	The Effects of Executives on Corporate Tax Avoidance	1163
McINNIS, JOHN	Earnings Smoothness, Average Returns, and Implied Cost of Equity Capital	315
McVAY, SARAH (and FENG)	Analysts' Incentives to Overweight Management Guidance When Revising Their Short-Term Earnings Forecasts	1617
MENON, KRISHNAGOPAL (and WILLIAMS)	Investor Reaction to Going Concern Audit Reports	2075

MILLER, BRIAN P.	The Effects of Reporting Complexity on Small and Large Investor Trading....2107
MILLS, LILLIAN F. (and BLOUIN and GLEASON and SIKES)	Pre-Empting Disclosure? Firms' Decisions Prior to FIN No. 48 791
MILLS, LILLIAN F. (and ROBINSON and SANSING)	FIN 48 and Tax Compliance 1721
NAGAR, VENKY (and CASKEY and PETACCHI)	Reporting Bias with an Audit Committee 447
NETESSINE, SERGUEI (and BAIMAN and SAOUMA)	Informativeness, Incentive Compensation, and the Choice of Inventory Buffer 1839
NEWMAN, D. PAUL (and LAUX)	Auditor Liability and Client Acceptance Decisions 261
NG, TERENCE BU-PEOW (and SHANKAR)	Effects of Technical Department's Advice, Quality Assessment Standards, and Client Justifications on Auditors' Propensity to Accept Client-Preferred Accounting Methods 1743
NORMAN, CAROLYN STRAND (and ROSE and ROSE)	Perceptions of Investment Risk Associated with Material Control Weakness Pervasiveness and Disclosure Detail....1787
ORMAZABAL, GAIZKA (and GOW and TAYLOR)	Correcting for Cross-Sectional and Time-Series Dependence in Accounting Research.... 483
PEECHER, MARK E. (and PIERCEY and RICH and TUBBS)	The Effects of a Supervisor's Active Intervention in Subordinates' Judgments, Directional Goals, and Perceived Technical Knowledge Advantage on Audit Team Judgments....1763
PETACCHI, PAOLO (and CASKEY and NAGAR)	Reporting Bias with an Audit Committee 447
PETERS, GARY F. (and MASLI and RICHARDSON and SANCHEZ)	Examining the Potential Benefits of Internal Control Monitoring Technology 1001
PETRONI, KATHY R. (and BAMBER and JIANG and WANG)	Comprehensive Income: Who's Afraid of Performance Reporting?.... 97
PIERCEY, M. DAVID (and PEECHER and RICH and TUBBS)	The Effects of a Supervisor's Active Intervention in Subordinates' Judgments, Directional Goals, and Perceived Technical Knowledge Advantage on Audit Team Judgments 1763

PIZZINI, MINA	Group-Based Compensation in Professional Service Firms: An Empirical Analysis of Medical Group Practices	343
POWNALL, GRACE (and BARTON and HANSEN)	Which Performance Measures Do Investors Around the World Value the Most—and Why?	753
QI, YAXUAN (and KIM)	Accruals Quality, Stock Returns, and Macroeconomic Conditions	937
RAJGOPAL, SHIVA (and BOWEN and CALL)	Whistle-Blowing: Target Firm Characteristics and Economic Consequences	1239
RAMESH, K. (and D'SOUZA and SHEN)	The Interdependence between Institutional Ownership and Information Dissemination by Data Aggregators	159
REFFETT, ANDREW B.	Can Identifying and Investigating Fraud Risks Increase Auditors' Liability?	2145
RESUTEK, ROBERT J.	Intangible Returns, Accruals, and Return Reversal: A Multiperiod Examination of the Accrual Anomaly	1347
RICH, JAY S. (and PEECHER and PIERCEY and TUBBS)	The Effects of a Supervisor's Active Intervention in Subordinates' Judgments, Directional Goals, and Perceived Technical Knowledge Advantage on Audit Team Judgments	1763
RICHARDSON, VERNON J. (and MASLI and PETERS and SANCHEZ)	Examining the Potential Benefits of Internal Control Monitoring Technology	1001
RIEDL, EDWARD J. (and ARMSTRONG and BARTH and JAGOLINZER)	Market Reaction to the Adoption of IFRS in Europe	31
ROBINSON, JOHN R. (and SIKES and WEAVER)	Performance Measurement of Corporate Tax Departments	1035
ROBINSON, LESLIE A. (and MILLS and SANSING)	FIN 48 and Tax Compliance	1721
ROBINSON, LESLIE A.	Do Firms Incur Costs to Avoid Reducing Pre-Tax Earnings? Evidence from the Accounting for Low-Income Housing Tax Credits	637
ROSE, ANNA M. (and ROSE and NORMAN)	Perceptions of Investment Risk Associated with Material Control Weakness Pervasiveness and Disclosure Detail	1787

ROSE, JACOB M. (and NORMAN and ROSE) — Perceptions of Investment Risk Associated with Material Control Weakness Pervasiveness and Disclosure Detail 1787

ŞABAC, FLORIN (and CHRISTENSEN and TIAN) — Ranking Performance Measures in Multi-Task Agencies 1545

SANCHEZ, JUAN MANUEL (and MASLI and PETERS and RICHARDSON) — Examining the Potential Benefits of Internal Control Monitoring Technology 1001

SANSING, RICHARD C. (and GUENTHER) — The Effect of Tax-Exempt Investors and Risk on Stock Ownership and Expected Returns 849

SANSING, RICHARD C. (and MILLS and ROBINSON) — FIN 48 and Tax Compliance 1721

SAOUMA, RICHARD (and BAIMAN and NETESSINE) — Informativeness, Incentive Compensation, and the Choice of Inventory Buffer 1839

SBARAGLIA, ANDREW M. (and BARUA and LIN) — Earnings Management Using Discontinued Operations 1485

SEYBERT, NICHOLAS — R&D Capitalization and Reputation-Driven Real Earnings Management 671

SHANKAR, PREMILA GOWRI (and NG) — Effects of Technical Department's Advice, Quality Assessment Standards, and Client Justifications on Auditors' Propensity to Accept Client-Preferred Accounting Methods 1743

SHAW, KENNETH W. (and ZHANG) — Is CEO Cash Compensation Punished for Poor Firm Performance? 1065

SHEN, MIN (and D'SOUZA and RAMESH) — The Interdependence between Institutional Ownership and Information Dissemination by Data Aggregators 159

SHIN, JAE YONG (and BOL and KEUNE and MATSUMURA) — Supervisor Discretion in Target Setting: An Empirical Investigation 1861

SIKES, STEPHANIE A. (and BLOUIN and GLEASON and MILLS) — Pre-Empting Disclosure? Firms' Decisions Prior to FIN No. 48 791

SIKES, STEPHANIE A. (and ROBINSON and WEAVER) — Performance Measurement of Corporate Tax Departments 1035

SISNEROS, CRAIG A. (and CREADY and LOPEZ) — The Persistence and Market Valuation of Recurring Nonrecurring Items 1577

SONG, CHANG JOON (and THOMAS and YI)	Value Relevance of FAS No. 157 Fair Value Hierarchy Information and the Impact of Corporate Governance Mechanisms 1375
STEFANIAK, CHAD M. (and HATFIELD and HOUSTON and USREY)	The Effect of Magnitude of Audit Difference and Prior Client Concessions on Negotiations of Proposed Adjustments 1647
STOCKEN, PHILLIP C. (and FISCHER)	Analyst Information Acquisition and Communication 1985
STUBBEN, STEPHEN R.	Discretionary Revenues as a Measure of Earnings Management 695
SUN, SUNNY Y. J. (and CHEN and WU)	Client Importance, Institutional Improvements, and Audit Quality in China: An Office and Individual Auditor Level Analysis 127
SUN, YAN	Do MD&A Disclosures Help Users Interpret Disproportionate Inventory Increases? 1411
TAN, HUN-TONG (and JAMAL)	Joint Effects of Principles-Based versus Rules-Based Standards and Auditor Type in Constraining Financial Managers' Aggressive Reporting 1325
TAYLER, WILLIAM B.	The Balanced Scorecard as a Strategy-Evaluation Tool: The Effects of Implementation Involvement and a Causal-Chain Focus 1095
TAYLOR, DANIEL J. (and GOW and ORMAZABAL)	Correcting for Cross-Sectional and Time-Series Dependence in Accounting Research 483
THOMAS, WAYNE B. (and FAN and BARUA and CREADY)	Managing Earnings Using Classification Shifting: Evidence from Quarterly Special Items 1303
THOMAS, WAYNE B. (and SONG and YI)	Value Relevance of FAS No. 157 Fair Value Hierarchy Information and the Impact of Corporate Governance Mechanisms 1375
TIAN, JIE (JOYCE) (and CHRISTENSEN and ŞABAC)	Ranking Performance Measures in Multi-Task Agencies 1545
TUBBS, RICHARD M. (and PEECHER and PIERCEY and RICH)	The Effects of a Supervisor's Active Intervention in Subordinates' Judgments, Directional Goals, and Perceived Technical Knowledge Advantage on Audit Team Judgments 1763
USREY, SPENCER (and HATFIELD and HOUSTON and STEFANIAK)	The Effect of Magnitude of Audit Difference and Prior Client Concessions on Negotiations of Proposed Adjustments 1647

WANG, ISABEL YANYAN (and BAMBER and JIANG and PETRONI)	Comprehensive Income: Who's Afraid of Performance Reporting?	97
WANG, ISABEL YANYAN (and BAMBER and JIANG)	What's My Style? The Influence of Top Managers on Voluntary Corporate Financial Disclosure	1131
WATTS, SUSAN G. (and BAGNOLI)	Oligopoly, Disclosure, and Earnings Management	1191
WEAVER, CONNIE D. (and ROBINSON and SIKES)	Performance Measurement of Corporate Tax Departments	1035
WEBER, JOSEPH (and BEATTY and LIAO)	Financial Reporting Quality, Private Information, Monitoring, and the Lease-versus-Buy Decision	1215
WEISS, DAN	Cost Behavior and Analysts' Earnings Forecasts	1441
WILLIAMS, DAVID D. (and MENON)	Investor Reaction to Going Concern Audit Reports	2075
WILLIAMS, MICHAEL G. (and HUGHES and LEVINE)	Influence of Capital Gains Tax Policy on Credibility of Unverified Disclosures	719
WILLIAMSON, MICHAEL G. (and KOONCE and WINCHEL)	Consensus Information and Nonprofessional Investors' Reaction to the Revelation of Estimate Inaccuracies	979
WILLIAMSON, MICHAEL G. (and KACHELMEIER)	Attracting Creativity: The Initial and Aggregate Effects of Contract Selection on Creativity-Weighted Productivity	1669
WINCHEL, JENNIFER (and KOONCE and WILLIAMSON)	Consensus Information and Nonprofessional Investors' Reaction to the Revelation of Estimate Inaccuracies	979
WU, DONGHUI (and CHEN and SUN)	Client Importance, Institutional Improvements, and Audit Quality in China: An Office and Individual Auditor Level Analysis	127
YEUNG, P. ERIC (and BAMBER and HUI)	Managers' EPS Forecasts: Nickeling and Diming the Market?	63
YI, HAN (and SONG and THOMAS)	Value Relevance of FAS No. 157 Fair Value Hierarchy Information and the Impact of Corporate Governance Mechanisms	1375
ZHANG, MAY H. (and SHAW)	Is CEO Cash Compensation Punished for Poor Firm Performance?	1065
ZHAO, YUPING (and KANG)	Information Content and Value Relevance of Depreciation: A Cross-Industry Analysis	227

BOOK REVIEWS

Author	*Title*	*Reviewer*	*Page*
Barth, Mary E.	Research, Standard Setting, and Global Financial Reporting	MARK H. LANG	381
Bloomfield, Robert and Rennekamp	Experimental Research in Financial Reporting: From the Laboratory to the Virtual World	KATHRYN KADOUS	1473
Christensen, Peter O. and Feltham	Equity Valuation	XIAO-JUN ZHANG	1809
Clikeman, Paul M.	Called to Account: Fourteen Financial Frauds that Shaped the American Accounting Profession	W. MORLEY LEMON	1811
Criscione, E. Richard	Abraham J. (Abe) Briloff: A Biography	GEORGE FOSTER	1814
De Bondt, Werner (editor)	Financial Accounting and Investment Management	STEVEN J. MONAHAN	1816
	Developments in New Reporting Models, Information for Better Markets	CHRISTINE A. BOTOSAN	2169
Easton, Peter	Estimating the Cost of Capital Implied by Market Prices and Accounting Data	CHARLES M. C. LEE	745
Francis, Jennifer, Olsson, and Schipper	Earnings Quality	MAUREEN McNICHOLS	382
Kanodia, Chandra	Accounting Disclosure and Real Effects	PIERRE JINGHONG LIANG	1119
Kaplan, Robert S. and Norton	The Execution Premium: Linking Strategy to Operations for Competitive Advantage	SHAHID ANSARI	1475
McPhail, Ken and Walters	Accounting and Business Ethics: An Introduction	JAMES C. GAA	1817
Melumad, Nahum D. and Nissim	Line-Item Analysis of Earnings Quality	SARAH E. McVAY	1121

Mosso, David	Early Warning and Quick Response: Accounting in the Twenty-First Century	ROBERT J. SACK	1122
Winkler, Othmar W.	Interpreting Economic and Social Data: A Foundation of Descriptive Statistics	THOMAS R. DYCKMAN	1820
Zeff, Stephen (Editor)	Principles before Standards: The ICAEW's 'N Series' of Recommendations on Accounting Principles 1942–1969	CHRISTOPHER NAPIER	1477
	Capsule Commentary	STEPHEN A. ZEFF	748, 1125

NOVEMBER 2010 PLACEMENT ADS

The deadline for free position ads to be included in this section is two months prior to the date of publication. *The Accounting Review* is published in January, March, May, July, September, and November. Position ads in the "Placement Ads" are free with the purchase of a job posting in the AAA Career Center. For more information on how to purchase a job posting in the Career Center and receive your free position ad in *The Accounting Review*, go to the AAA website at http://aaahq.org and click on "Career Center," or call our office at 941-921-7747.

601 SIENA COLLEGE, Accounting Department invites applications for a tenure-track position beginning September 2011. Applications will be accepted until position is filled. Rank and salary dependent upon experience and credentials. A Ph.D./D.B.A. with concentration in accounting is expected, relevant business experience desirable. An experienced tax professional with an L.L.M. is acceptable for someone desiring a position teaching tax. Research and professional development activities appropriate to academically qualified status for AACSB accreditation and the School of Business mission are necessary. If sufficient Academically Qualified individuals are not located, then individuals who are presently Professionally Qualified and willing to pursue a Ph.D. or D.B.A. may be considered for a position as a Visiting Professor. Individuals hired under such conditions would have their status evaluated annually and renewal would be dependent upon continued satisfactory progress toward a Ph.D. or D.B.A. Siena College is an independent liberal arts college in the Catholic and Franciscan tradition, located in New York's Capital District. The college enrolls 2,800 full-time students, 1,200 in the School of Business. Please forward applications to: Accounting Search Committee, c/o School of Business, Siena College 515 Loudon Road, Loudonville, NY 12211-1462. Siena College is an Equal Opportunity Employer and encourages applications from all qualified candidates.

602 BOISE STATE UNIVERSITY, Department of Accountancy invites applications for a tenure-track managerial accounting faculty position at either the Assistant or Associate rank starting Fall 2011. A Ph.D. in business with an accounting emphasis is required. Professional certification is preferred. Expectations include a commitment to high-quality teaching, scholarship leading to conference presentations and refereed publications, and university and community service. Both the College and the Department's undergraduate and graduate programs are AACSB-accredited. BSU is located in downtown Boise, Idaho, a thriving community offering city amenities and convenient access to an array of outdoor recreational opportunities. Salary is competitive within the Western region. Review of applications begins immediately and continues until the position is filled. For more information regarding this position and application procedures, please visit: http://hrs.boisestate.edu/joblistings/faculty/; or contact: Dr. Paul Bahnson at pbahnson@boisestate.edu. Boise State University is an EEO/AA institution, veteran's preference.

603 DRAKE UNIVERSITY, College of Business and Public Administration, School of Accounting, invites applications for an Assistant/Associate Professor of Accounting position, starting in Fall 2011, teaching primarily cost/managerial accounting at the undergraduate and graduate level. The normal teaching requirements are nine credit hours per semester. Demonstrated effectiveness in teaching and continuing professional development is expected. A doctorate in Accounting at time of appointment is preferred, but ABD candidates close to completion will be considered. Candidates must also be able to provide evidence of demonstrated teaching effectiveness at the university-level and continuing professional development. Candidates for the Associate level must show a record of scholarly publications and an active research stream. Salary is competitive with peer AACSB institutions. Drake University is a private, comprehensive university dedicated to quality teaching. It is located in Des Moines, Iowa, an insurance and financial center with a metropolitan population over 500,000. Enrollment is selective and limited to approximately 5,500 students in the schools and colleges of the University. Both the College of Business and Public Administration and School of Accounting hold AACSB accreditation. Send application letter with current vita listing three (3) references to: Dr. Stephen Gara, Search Committee Chair, School of Accounting, Drake University, 2507 University Avenue, Des Moines, IA 50311; Email: Stephen.Gara@drake.edu; URL: http://www.drake.edu/hr/. Drake University is an EEO.

604 UNIVERSITY OF DAYTON, Department of Accounting invites applications for a tenure-track position at the Assistant Professor level beginning August 16, 2011. The ability to teach introductory and advanced courses in U.S. taxation at the undergraduate and graduate levels, as well as other related accounting or law courses, is required, as is previous successful professional tax experience, CPA certification, and tax teaching experience. Candidates must meet, and have the ability to maintain AACSB business and accounting accreditation requirements for "Academically Qualified" (AQ) status. This includes a Ph.D. from an AACSB-accredited institution or a combination of a master's degree in business with an emphasis in accounting, a J.D., and a L.L.M. degree with specialization in taxation. Candidates should have the potential for appropriate teaching, research, and service. We prefer candidates who have experience teaching students from diverse racial and ethnic backgrounds and who have demonstrated interest in mentoring traditionally underrepresented students in their area of expertise. We prefer a candidate who will embrace values consistent with Marianist and Catholic traditions. For more information on UD, see: http://www.udayton.edu/business/accounting/index.php. The complete position description and the process for applicants to apply are available at: https://jobs.udayton.edu/. Closing Date: September 30 or until filled. UD is Ohio's largest independent university and one of the nation's largest Catholic universities. UD is firmly committed to the principle of diversity and is an Affirmative Action/Equal Opportunity Employer. Persons of color, women, individuals with disabilities, and veterans are encouraged to apply.

605 THE UNIVERSITY OF NORTH CAROLINA AT PEMBROKE, Department of Accounting and Information Technology, School of Business invites applications for one tenure-track position in accounting to begin January 2011 or later. Rank and salary depend upon qualifications. Candidates with a Ph.D. in Accounting are preferred. ABDs will be considered if degree is to be completed by the hiring date. Applicants from all areas of accounting will be considered. Professional certifications and experience are preferred. Applicants must complete the online application form through UNCP's Online Employment System found at http://www.uncp.edu/hr. UNC Pembroke is an Equal Employment/Affirmative Action Employer.

606 WILFRID LAURIER UNIVERSITY, School of Business and Economics invites applications for one position as Assistant and one position as Associate or Full Professor in the Accounting area, commencing July 1, 2011. The positions are subject to budgetary approval. Preference will be given to applicants with research and teaching interests in Auditing, Financial Accounting, and Management Accounting. Having a professional designation will be an asset. Applicants without a designation will be encouraged to obtain one after they are hired. Applicants for these positions must meet the minimum qualification of a Ph.D. in Accounting or ABD. Applicants must also exhibit an interest in and the capacity to conduct quality scholarly research and exhibit the potential for a high level of teaching competence at both the graduate and undergraduate levels. For the senior position, preference will be given to applicants with past Ph.D. student supervision experience and an established record of research publications in accounting. All qualified candidates are encouraged to apply; however, Canadians and Permanent Residents will be given priority. Applicants should send a CV, a summary of research objectives and teaching experience, a sample research article, and the names of three references. Electronic applications are encouraged. Direct your application and supporting documents by November 30, 2010 to: Dr. Peter Carayannopoulos, Associate Dean of Business, Faculty Development & Research, School of Business and Economics, Wilfrid Laurier University, 75 University Avenue, West Waterloo, Ontario N2L 3C5, CANADA; Email: bizrecruit@wlu.ca. Wilfrid Laurier University is committed to equity and values diversity. We welcome applications from qualified individuals of all genders and sexual orientations, persons with disabilities, Aboriginal persons, and persons of a visible minority. Members of the designated groups must self-identify to be considered for employment equity.

607 OHIO NORTHERN UNIVERSITY seeks applicants. Responsibilities: Administer all aspects of the Accounting graduate program including admissions. Primary teaching responsibilities, under a two-two teaching load, will include Auditing and Accounting Information Systems and other courses consistent with the candidate's professional interests and background. A well-established record of scholarship is expected. Qualifications: Doctorate in Accounting and demonstrated commitment to excellence in teaching (both at the undergraduate and graduate levels) and academic advising. Evidence of research publications in peer-reviewed journals that would qualify the applicant for rank of Associate Professor. Teaching experience involving distance education a plus. Compensation: Salary dependent upon experience and qualifications. Fringe benefits include TIAA-CREF retirement program, medical insurance, disability benefits, life insurance, educational benefits, and tuition remission for dependents attending Ohio Northern University for an initial degree. Terms of Employment: Tenure-track nine-month appointment. Employment Date: September 6, 2011. Please contact Dr. James W. Fenton via email at: j-fenton.1@onu.edu; or phone: (419) 772-2070.

608 UNIVERSITY OF SAN FRANCISCO, School of Business and Professional Studies seeks applicants for one tenure-track faculty position in the field of accounting at the rank of Assistant or Associate Professor starting Fall 2011. The School seeks applicants conducting research in all aspects of accounting but gives preference to either of the following areas: (1) Auditing focusing on professional ethics, behavioral aspects of auditing, forensic accounting, fraud, or other aspects of auditing and assurance services; (2) Financial Accounting focusing on international harmonization and IFRS, consolidations, or other aspects of financial accounting and reporting. A successful candidate must meet the following minimum qualifications: (1) an earned doctorate in accounting or a related field by the time of appointment; (2) an expectation of excellence in teaching; and (3) an established or emerging research program. The successful candidate will teach both required and elective undergraduate and M.B.A. courses. You will undertake scholarly research in a collaborative and supportive environment that welcomes inter-disciplinary and practice-oriented scholarship as well as discipline-based and theoretical work. The School of Business and Professional Studies offers bachelor's and master's degree programs. The School is housed in Malloy Hall, a new, purpose-built facility with fully modern classrooms, offices, and informal workspaces. Further information regarding USF and the School of Business and Professional Studies is available at: http://web.usfca.edu/bps.

609 UNIVERSITY OF SOUTH FLORIDA, ST. PETERSBURG seeks applicants for an Assistant, Associate, or Full Professor in Accounting. Responsibilities will include teaching undergraduate courses and/or graduate courses in accounting; participating in college and university committees; leading through professional and community service. Research and publication will be expected consistent with rank. SALARY $90,000 and higher, depending on faculty credentials and budgetary approvals. Minimum Qualifications: Doctoral degree in accounting from an AACSB-accredited institution. Preferred Qualifications: Candidates possessing the following additional qualifications will be given preference: (1) primary teaching and research interests in the intermediate financial accounting and advanced financial accounting topics. A secondary interest in forensic accounting and fraud examination is desirable; (2) online and/or graduate teaching experience; (3) published research in accounting journals; (4) recent experience (past five years) teaching upper-level accounting at an AACSB-accredited institution; and (5) Research or teaching in corporate social responsibility (CSR) topics. Please contact John F. Jewell at email: jewell@mail.usf.edu; or phone: (727) 873-4110. Applicants should apply for this position by visiting: https://employment.usf.edu/applicants/jsp/shared/search/SearchResults_css.jsp

610 UNIVERSITY OF NOTRE DAME, Department of Accountancy, invites applications for position openings at all levels, Assistant, Associate, or Full Professor. All research areas will be considered. The appointment is expected to begin July 1, 2011. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please submit curriculum vita, three letters of recommendation, teaching evaluations, and an example of scholarly work to: acctdept@nd.edu. Also, address inquiries to: acctdept@nd.edu. The deadline for applications is January 15, 2011. For more information about the University of Notre Dame, the Mendoza College of Business, and the Department of Accountancy please visit the department's website at: http://business.nd.edu/accountancy/. The University of Notre Dame is an Affirmative Action Employer with a strong commitment to fostering a culturally diverse atmosphere for faculty, staff, and students. Women, minorities, and those attracted to a university with a Catholic identity are especially encouraged to apply.

611 GONZAGA UNIVERSITY seeks applicants for an Assistant or Associate Professor, with strong emphasis in taxation or audit/assurance. Required: Ph.D. in accounting or closely related discipline, at or near completion, or J.D.; demonstrated evidence of teaching effectiveness and scholarly research. Desired: Professional certification and industry/public accounting experience Salary: Commensurate with experience, AACSB-competitive. Gonzaga University is a private liberal arts institution of approximately 6,000 students, recognized by U.S. News & World Report as among the top regional universities in the West. There are approximately 1,100 undergraduate and 250 graduate students enrolled in the business school, (about 25–30 percent are in the accounting program). Gonzaga accounting students have consistently performed well on the CPA exam, achieving a first-time pass rate well in excess of the national average, and are heavily recruited by CPA firms and industry. Gonzaga's educational mission is set in the Jesuit tradition with emphasis on the moral and ethical implications of learning, living, and working within today's global societies. The school is AACSB-accredited and offers B.B.A., M.B.A., and M.Acc. degrees. In addition, the accounting program recently achieved separate accreditation. The School is housed in a technologically up-to-date 60,000 sq ft building located along the Spokane River, in the midst of Gonzaga's beautifully landscaped 115-acre campus. Gonzaga University is a Jesuit, Catholic, humanistic institution, and is interested in candidates who will contribute to its distinctive mission. Gonzaga University is an AA/EEO employer committed to diversity. Send applications to: Dr. Gary Weber at: weber@gonzaga.edu.

612 MARQUETTE UNIVERSITY invites applications for one tenure-track position in Accounting Information Systems beginning Fall 2011. Applicants should hold a doctoral degree or be in the final stages of dissertation. The position will be filled at the rank of Assistant Professor. The primary teaching responsibility will be in Accounting Information Systems at the undergraduate and graduate levels with a secondary area in Managerial, Fraud, Audit, Financial, or International. The normal teaching load is 12 semester hours (6 hours each semester) during the first year and 15 semester hours thereafter. Salary and research support are competitive. The department has 12 full-time faculty. Marquette offers AACSB-accredited undergraduate and graduate degrees in Accounting and Business. All applicants must apply online at: http://careers.marquette.edu/applicants/Central?quickFind=52728. Inquiries may be made to: Michael D. Akers, Professor and Chair, Department of Accounting, PO Box 1881, College of Business Administration, Marquette University, Milwaukee, WI 53201-1881; Email: Michael.Akers@Marquette.edu. Marquette University is an Affirmative Action/Equal Opportunity Employer.

613 BROCK UNIVERSITY invites applications for one (1) probationary tenure-track position in Accounting starting July 1, 2011. The position rank is open depending upon qualifications. The availability of this position is subject to final budgetary approval. Job Qualifications: A Ph.D. in Accounting is required (completed or near completion). Previous academic experience and professional designation are desirable, but not required. We seek candidates with the potential for excellence in teaching and research. Responsibilities: This is a probationary tenure-track position requiring quality research, teaching, and service. We welcome applications from candidates with interest in any area of accounting research and teaching. The Department of Accounting has 25 full-time faculty members and is in the Faculty of Business, which has approximately 90 full-time members. We offer a full range of graduate and undergraduate programs, both research and professionally oriented: Master of Accountancy, Master of Science with an Accounting stream, M.B.A. with an optional Accounting stream, Bachelor of Accounting, and Bachelor of Business Administration with an optional Accounting stream. Co-op options are available in both undergraduate and graduate programs. The Faculty of Business is AACSB-accredited. Our accounting programs are accredited by the Institute of Chartered Accountants of Ontario, the Society of Management Accountants of Ontario, and the Certified General Accountants of Ontario. More information on Brock University can be found on the University's website: http://www.brocku.ca.

614 BINGHAMTON UNIVERSITY, SUNY, AACSB-accredited School of Management invites applications for an Assistant Professor position in Accounting for Fall 2011, pending budget approval. Candidates must have (or expect) a Ph.D. degree by Fall 2011. High-quality research as evidenced by publication in the top-tier accounting journals is expected. Research is supported with a two-course/semester teaching assignment. The School of Management also supports research with subscriptions to WRDS and other standard databases. The University offers bachelor's, master's, and Ph.D. programs in accounting. More details are available at: http://som.binghamton.edu/faculty/index.htm. Located three hours from NYC and one hour from Syracuse and Ithaca, Binghamton is a low-cost-of-living area with 200,000 people. Interested candidates should post a curriculum vita and other supporting documents to the appropriate position listing at: http://binghamton.interviewexchange.com/. Application materials to be submitted electronically. Candidates must have (or expect) a Ph.D. degree by Fall 2011. Binghamton University is an Equal Opportunity/Affirmative Action Employer.

615 PURDUE UNIVERSITY, Krannert School of Management invites applicants for one or more tenure-track appointments in accounting beginning Fall 2011. Rank is at either the Assistant or Associate Professor level. Candidates for an Assistant Professor position should have a doctorate in accounting, management, or closely related field, and exhibit the capacity for research publishable in top-tier accounting journals and excellence in teaching. ABDs will be considered. Candidates for an Associate Professor position should have an earned doctorate and an established record of excellence in research, teaching, and service on doctoral dissertation committees. A tenured Associate Professor appointment is possible for a candidate with the appropriate credentials. The successful candidate will teach financial or managerial accounting to undergraduate and/or graduate students, conduct research in their field of expertise, and participate in other school activities. Opportunities to teach in the executive program may also be available to candidates with prior teaching experience. We offer a competitive compensation package. Applications will be reviewed beginning October 15, 2010. Applications should include a CV, representative research output, and evidence of teaching ability. Applicants should also be prepared to provide at least three letters of reference. Applications should be sent to: Professor Mark Bagnoli, Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907. Purdue University is an Equal Opportunity/Equal Access/ Affirmative Action Employer fully committed to achieving a diverse workforce.

616 MASSACHUSETTS INSTITUTE OF TECHNOLOGY, Sloan School of Management, Accounting Group is seeking new faculty at all levels (Assistant, Associate, and Full Professor) for the academic year beginning July 1, 2011, or as soon as possible thereafter. Applicants will be evaluated on their potential for and/or record of excellence in research and teaching; faculty responsibilities include teaching financial accounting at graduate and undergraduate level. Please include evidence of research such as a publication, working paper, or dissertation proposal with other application materials. The deadline for applications is November 30, 2010. However, applications will be evaluated as soon as they are complete, if submitted at an earlier date. Applications must include a cover letter, curriculum vitae, and letters of recommendation. Applications should be sent to: Professor Michelle Hanlon Chair, Accounting Group Search Committee, MIT Sloan School of Management, 77 Massachusetts Avenue, E62-668, Cambridge, MA 02139. For overnight deliveries, please use the following address: MIT Sloan School of Management, 100 Main Street, E62-668, Cambridge, MA 02142. MIT is an Equal Opportunity Employer committed to building a culturally diverse intellectual community and strongly encourages applications from women and minorities.

617 INDIANA UNIVERSITY, Kelley School of Business, Department of Accounting invites applications for tenure-track/tenured faculty positions effective Fall 2011. Preference will be given to those with teaching and research interests in Auditing. Compensation levels are competitive. Candidates must possess a doctoral degree by Fall 2011 and have a strong commitment to excellence in scholarly research and teaching. Interested candidates should send a current vita, a research paper suitable for presentation at a workshop, three letters of reference (six letters for tenured offers), and evidence of teaching ability to the contact person indicated below. Applications received before November 30, 2010 will be given full consideration. Contact address: Laureen Maines, Chair, Department of Accounting, Kelley School of Business, 1309 E. 10th Street, Suite 540, Bloomington, IN 47405-1701; Email: lmaines@indiana.edu. Indiana University is an Affirmative Action/Equal Opportunity Employer committed to excellence through diversity. The University actively encourages applications of women, minorities, and persons with disabilities.

618 UNIVERSITY OF HAWAI'I AT WEST O'AHU, Professional Studies seeks candidates for Assistant/Associate Professor in Accounting, position number 85439, tenure-track, full-time, nine-month, to begin August 1, 2011. Duties: Assistant Professor. This position is responsible for teaching accounting courses in the Business Administration program. Responsibilities include, but are not limited to: teaching courses in Accounting (primary need is Auditing); developing new courses; providing academic advising, and supervising practicum students; and teaching distance education courses. The successful candidate is also expected to conduct research or applied scholarship in the discipline and to participate in faculty governance. Associate Professor: Same as above. Minimum qualifications for Assistant Professor: Earned doctorate degree from an accredited college or university in Accounting, or a related field with an emphasis in Accounting. (ABD candidates are eligible to apply but must complete all degree requirements prior to appointment.) Evidence of high-quality teaching and potential for conducting relevant scholarly and applied research. Teach distance education courses. Minimum qualifications for Associate Professor: All of the above plus five years of teaching experience. Evidence of publications in peer-reviewed journals in the area of Accounting. Desirable qualifications for Assistant Professor: Ability to teach Auditing, Intermediate Financial Accounting, Advanced Accounting, and/or Governmental Accounting. Ability to interact effectively in a multicultural setting, experience teaching diverse populations. Demonstrated ability to teach distance education courses. CPA or other certification. Desirable qualifications for Associate Professor: Ability to teach Auditing, Intermediate Financial Accounting, Advanced Accounting, and/or Governmental Accounting. Ability to interact effectively in a multicultural setting, experience teaching diverse populations. Demonstrated ability to teach distance education courses. CPA or other certification. Pay range: The salary will be commensurate with qualifications and experience. Applicants may apply to: http://www.uhwo.hawaii.edu/vacancies

619 UNIVERSITY OF INDIANAPOLIS, School of Business is seeking applications for an Assistant Professor of Accounting. This is a tenure-track position with a start date of the Fall 2011. The successful candidate will participate in teaching, service, and scholarship as it relates to the field of accounting. The ideal candidate will have a strong background in teaching undergraduate accounting, as well as be professionally qualified. Ph.D./D.B.A. in accounting is preferred (ABD candidates may apply). An M.B.A./CPA is also acceptable, given extensive professional and academic experience. Candidates must have demonstrated proficiency in teaching and knowledge of the accounting field. Applied scholarship and active participation in the discipline is a plus. The University of Indianapolis, School of Business is nationally accredited by the Association of Collegiate Business Schools and Programs (ACBSP). Applications may be submitted at: https://jobs.uindy.edu/.

620 THE UNIVERSITY OF TENNESSEE AT KNOXVILLE, Department of Accounting and Information Management is accepting applications for a full-time, tenure-track position to begin August 1, 2011. Candidates should hold a doctoral degree with an emphasis in accounting or expect to complete all degree requirements by his or her starting date. The ideal candidate will have a research interest in corporate governance and a teaching interest in auditing or financial accounting. Candidates for Assistant Professor should demonstrate excellent teaching potential and have an academic background that strongly suggests an ability to establish an ongoing program of substantive research. Candidates for advanced rank should demonstrate excellent achievement in teaching and have a substantive research record that reflects a significant ongoing research program. All faculty searches are subject to the availability of funding. Please send cover letter, vita, and references, preferably by email, to: Kenneth E. Anderson, Search Committee Chair, Department of Accounting and Information Management, CBA, University of Tennessee Knoxville, Tennessee 37996-0560; Email: kea@utk.edu. The University of Tennessee is an EEO/AA/Title VI/Title IX/Section 504/ADA/ADEA institution in the provision of its education and employment programs and services. All qualified applicants will receive equal consideration for employment without regard to race, color, national origin, religion, sex, pregnancy, marital status, sexual orientation, gender identity, age, physical or mental disability, or covered veteran status.

621 EASTERN MICHIGAN UNIVERSITY seeks applicants for one tenure-track position at the Assistant/Associate level in accounting beginning Fall 2011, subject to funding approval. All areas of specialization will be considered. Individuals who have a Ph.D. in hand are strongly preferred, but, at minimum, an ABD from an AACSB-accredited university or its equivalent leading to a Ph.D. is required. Professional certifications such as a CPA/CMA/CIA and/or business experience are desirable. Salary is competitive. Faculty members are expected to teach specific courses, conduct research, and participate in university functions. EMU is a public, comprehensive university of 23,000 students located in southeastern Michigan and is an Equal Opportunity employer. The College of Business is AACSB-accredited. To apply for Posting #fa1101e go to: https://www.emujobs.com/, click on the "View/Apply for Faculty and Administrative Positions" button located in the middle of the page. If you have questions, contact Professor Linda Burilovich, Chair of the Personnel Committee at: linda.burilovich@emich.edu. Review of applications will continue until the position is filled.

622 IDAHO STATE UNIVERSITY, Department of Accounting seeks a full-time tenure-track Assistant or Associate Professor for a position available for the 2011–2012 academic year. Salary will be commensurate with experience and qualifications. Both the College of Business and the Accounting Program are accredited by AACSB. Our accounting faculty support excellent undergraduate and M.Acc. programs. Responsibilities: Teaching at the undergraduate and graduate levels with preferred emphasis areas in financial accounting. A primary interest in auditing with a secondary interest in financial accounting will also be considered. Normal course load includes two preps and three sections. Peer-reviewed scholarly work and service are expected. Minimum Qualifications: The successful candidate will demonstrate a collegial attitude, excellent communication skills, and an eagerness to support our students through a mentoring/ advising process. Contact Information: To apply, send a letter of interest; resume; and the names, addresses, and telephone numbers of three professional references to: Dr. Bob Picard, Chair, Department of Accounting, College of Business, Idaho State University, Campus Box 8020, Pocatello, ID 83209-8020. Idaho State University is an AA/EOE. Minorities, women, and veterans are encouraged to apply.

623 FLORIDA INTERNATIONAL UNIVERSITY, School of Accounting invites applications for instructors, clinical or tenure-track positions at all ranks in all areas of accounting and for instructors of business law beginning Fall 2009. Doctorate or L.L.M. (Taxation) required for accounting positions; J.D. required for business law positions. Preference may be given to applicants with publications in their respective fields. Expectations include strong commitment to excellence in graduate and undergraduate teaching, research, and service. Duties may include teaching in weekend and overseas programs. The School of Accounting is separately accredited by AACSB and offers B.Acc., M.Acc., M.S.T. and Ph.D. programs. FIU is identified in the Carnegie classification as a Research University/High Research Activity institution with enrollments in excess of 38,000 students. FIU is a growing public university that admitted its first class in 1972. It has a diverse student body including both full-time residential students as well as commuters. FIU's modern picturesque campus with tropical flora is located in the suburbs of Miami. To apply, send a vita and names of five references to: Ana Estevill, Florida International University, School of Accounting, 11200 SW 8th St., RB 210, University Park, Miami, FL 33199; Email: estevill@fiu.edu. FIU is an Equal Opportunity Employer. Applications from minority candidates are encouraged.

624 OKLAHOMA STATE UNIVERSITY, Spears School of Business, School of Accounting invites applications beginning Fall 2011 at the Assistant, Associate, or Full Professor level. All areas of research and teaching emphasis will be considered. We are especially interested in faculty with a proven research record indicating the ability to publish in "A" level journals. Applicants must have an appropriate earned doctorate and those at the senior level should have a distinguished ongoing record of scholarship. All applicants must have a commitment to instructional excellence. The position is located on the OSU/Stillwater campus, but could be called upon to occasionally teach on the OSU/Tulsa campus. Applicant screening will begin immediately and will continue until the position is filled. We intend to hold campus interviews for the position during the Fall 2010 semester. All positions are contingent upon available funding. Interviews for a second open position (pending funding approval) will be held during the Spring 2011 semester. A letter of application, curriculum vitae, and a summary of teaching evaluations should be sent to: Don Herrmann, Head, School of Accounting, Oklahoma State University, 401 Business Building, Stillwater, OK 74078; Phone: (405) 744-5123; Email: osu-soa@okstate.edu. Oklahoma State University is an Affirmative Action/Equal Opportunity/E-Verify employer committed to diversity. Please contact us for details.

625 FAIRLEIGH DICKENSON UNIVERSITY, Department of Accounting, Taxation, and Law, Silberman College of Business is accepting applications for a tenure-track position in Accounting at the Assistant/Associate Professor rank. Silberman College has AACSB-accredited undergraduate and graduate programs. Candidates should have a Ph.D. in Accounting from an AACSB-accredited institution (or ABD nearing completion of dissertation) with evidence of or potential for high-quality teaching and research. The successful candidate will be expected to teach undergraduate and graduate level accounting courses. The department offers an undergraduate in accounting, M.S. Accounting, and M.S. Taxation, with over 300 majors enrolled. The normal teaching load is 9 credit hours per semester. Salary and benefits are competitive with AACSB-accredited institutions. Employment is contingent upon a satisfactory background check. Candidates for hire will be required to sign a waiver authorizing the background check and produce a Social Security Card.

626 TEXAS A&M UNIVERSITY–CORPUS CHRISTI, College of Business, Department of Accounting and Business Law invites applications for a tenure-track Assistant Professor position beginning as early as Fall 2010, subject to final budgetary approval. The position is open to all areas of accounting specialization. Applicants should possess a doctoral degree in accounting from an AACSB-accredited institution and demonstrate a strong commitment to scholarly research and teaching excellence. Certification and professional experience are preferred. Salary is competitive. The College of Business undergraduate and graduate business and accounting programs are accredited by AACSB International and the College will be moving into a new building early in the Spring 2011. Please submit applications online at: http://islanderjobs@tamucc.edu. For more information on the position, the University, and the College go to: http://www.cob.tamucc.edu. TAMUCC is an AA/EEO employer committed to excellence (http://www.tamucc.edu/~hrweb/).

627 EASTERN ILLINOIS UNIVERSITY invites applicants for a 9-month tenure-track position at the Assistant or Associate level in Accountancy. Position begins Fall 2011. Candidate should possess expertise and teaching experience in either Financial or Tax Accounting, although other specialties will be considered. The successful candidate will also conduct research in accountancy and participate in curriculum development and appropriate committees. A Ph.D. or D.B.A. is preferred, although ABDs close to completion will be considered. Candidates should demonstrate a teaching, research, and service record commensurate with the desired professional rank. Evidence of strong instructional effectiveness and scholarly productivity is essential, as are strong communication and interpersonal skills. Business experience and professional certification are desirable. Applicants should submit electronically the following materials: (1) application letter, (2) curriculum vitae, and (3) names and contact information for three references to: Dr. Hank Davis, Director of Accountancy, at: busrecrt@eiu.edu with "Accounting Position Application" in the subject line. Finalists will be asked to provide transcripts. Review of applications will begin on August 17, 2010.

628 UNIVERSITY AT ALBANY, SUNY, Department of Accounting and Law invites applications for one tenure-track position at the Assistant Professor level starting January 1, 2010, subject to budgetary approval and available funds. The candidates must have a doctoral degree from a college or university accredited by a U.S. Department of Education (DOE) or internationally recognized accrediting organization and teaching experience in managerial/cost/financial accounting. Research interests in empirical managerial/financial accounting are preferred. The successful candidate will demonstrate evidence of a strong commitment to scholarly research and excellence in teaching, and an ability to work with and/or instruct diverse groups of people is required. Preference will be given to candidates with a previous record of success at the Assistant Professor level. Candidates are requested to submit their vitae and contacts for three references by email to: Ingrid E. Fisher, Accounting Search Chair, Department of Accounting and Law, School of Business, The University at Albany, State University of New York; Email: ifisher@uamail.albany.edu. Review of applications will begin immediately and continue until the position is filled. The University at Albany is an EEO/AA/IRCA/ADA employer.

629 UNIVERSITY OF OREGON, Lundquist College of Business University, Department of Accounting is seeking to fill one tenure-track position in accounting, subject to funding availability. Rank: Assistant Professor. Employment Beginning: September 16, 2011. Description: The Lundquist College is committed to publishing research in top accounting journals, providing a high level of service to students and interacting with the professional community. Responsibilities include teaching accounting courses at the undergraduate and graduate levels as well as supervising Ph.D. students. Promotion and tenure depend on the quality of research and teaching. The successful candidate will have the ability to work effectively with faculty, staff, and students from diverse backgrounds. Qualifications: Completed or nearly completed Ph.D. degree with demonstrated potential for teaching excellence and the ability to publish economics-based research in high-quality academic journals. Salary: Competitive. Application Procedure: Send letter of interest, curriculum vitae, copies of research papers, evidence of teaching performance, and three letters of reference to: ActgOpenSearch@lcbmail.uoregon.edu. Application Deadline: Review of applications begins December 1, 2010 and will continue until a suitable candidate is identified. The University of Oregon is an Equal Opportunity/Affirmative Action institution committed to cultural diversity and compliance with the Americans with Disabilities Act.

630 UNIVERSITY OF SOUTHERN CALIFORNIA, Marshall School of Business, Leventhal School of Accounting seeks a Dean. The Leventhal School, named in honor of Kenneth and Elaine Leventhal, who have provided a significant endowment for the School, offers undergraduate and graduate programs in accounting that are routinely ranked among the top 10 programs in accounting. The School's faculty members are thought-leaders in the practice and scholarship of accounting. The Leventhal School is committed to excellence in teaching and research. The Marshall School of Business, of which the Leventhal School is a part, is ranked in the top-tier business schools in undergraduate and graduate education, and is a leading producer of scholarly research in all fields of business. Candidates for the position of Dean should be established scholars in accounting or accounting professionals with a significant record of management and an understanding of the environment of a major research university. The successful candidate will be able to advance the School's thought leadership, cultivate and strengthen the School's educational programs, work effectively as part of the Marshall School, and coordinate development efforts with the Marshall School and the University of Southern California. Nominations and applications should be directed to: Chair, Leventhal School Dean Search Committee Marshall School of Business University of Southern California 3670 Trousdale Parkway, BRI 101, Los Angeles, CA 90089-0802. Located near the heart of Los Angeles, the University of Southern California is home to approximately 31,000 students and 4,300 faculty. USC Marshall is renowned for its high-ranking undergraduate, graduate, international, and executive education programs, an exceptional faculty engaged in leading-edge research, a diverse and creative student body, and a commitment to technological advancement. The research productivity of Marshall's 200 full-time faculty ranks among the top 15 business schools in the world. The University of Southern California values diversity and is committed to Equal Opportunity in employment. Women and men, and members of all racial and ethnic groups, are encouraged to apply.

631 NORTHEASTERN UNIVERSITY is seeking applications for tenure-track positions effective September 2011 at the Assistant or Associate Professor rank. We seek candidates who have or expect to have a doctorate in accounting or a closely related business field before the start date of September 2011 and who are committed to excellence in research and teaching. A strong research program focused on top-tier accounting journals is desired and teaching experience in accounting is preferred. While the group needs faculty to teach in all areas, it has particular needs for faculty to teach all levels of financial and managerial accounting. The accounting group provides a supportive environment in which collaboration in research and teaching is encouraged. The position offers competitive compensation and teaching responsibilities that enable a significant level of high-impact research. Rank and salary are dependent on qualifications. Summer funding, graduate assistants, and travel support are available. Accounting courses are offered at both the undergraduate and at the graduate levels (M.B.A., M.S./M.B.A., M.S.A., and M.S. Tax). Please submit your application online at: http://www.cba.neu.edu/faculty/. Click on "Faculty Positions." Include a cover letter, your current CV, and two samples of your scholarly work. Questions regarding this position may be directed to: Dr. Marjorie Platt at: m.platt@neu.edu. Northeastern University is an Affirmative Action/Equal Opportunity/Title IX Employer and welcomes applications from minorities, women, and disabled persons.

632 MIAMI UNIVERSITY, Department of Accountancy invites applicants for a non-tenure-track position at the clinical faculty level beginning fall semester 2011. This position primarily is to teach and provide service to the department. At minimum, applicants must have an accounting-related master's degree and university-level teaching experience. A doctorate, professional certification, and relevant professional work experience are desirable. Salary and benefits are competitive. Both the undergraduate and graduate accounting programs are AACSB-accredited. Faculty must be either academically or professionally qualified. Information about the faculty and curriculum is available on the web at: http://www.sba.muohio.edu. Please submit your resume and letters from three references to: Dr. Marc A. Rubin, Department of Accountancy, Farmer School of Business, Miami University, Oxford, OH 45056; Email: rubinma@muohio.edu. Screening of applications begins October 1, 2010 and will continue until the position is filled. Campus crime and safety report: http://www.muohio.edu/righttoknow (hard copy upon request). Miami University is an EOE/AA employer with smoke-free campuses.

633 MAYO CLINIC, in Rochester, MN, is seeking an experienced Senior Financial Auditor to assist Audit Managers in the performance of team-based operational, compliance, financial statements, integrated and system reviews, as well as Enterprise Risk Management (ERM) implementation. In addition to your ability to lead audit projects to meet department metrics, the position calls for excellent communication and group-facilitation skills. To qualify, you must have three-to-five years of auditing experience (public accounting with a national or regional firm preferred). You must have current CPA, CIA, or CMA certification (or willingness to work toward certification) and familiarity with MS Office Suite; ACL and Auto Audit a plus. Expect to travel 20–25 percent of the time. To apply or learn more about this and other opportunities, please visit http://www.mayo-clinic-jobs.com/job/Financial-Operational-Sr_-Auditor-27782-Job/867056/.

634 UCLA Faculty Practice Group seeks accountant, responsible for providing general accounting assistance to the Account Manager. Primary responsibilities include processing and coordinating general ledger month-end closing with internal and external data customers and suppliers, financial reporting, and serving as the department liaison. The Benefits of Belonging: As a valued member of our staff, you'll enjoy outstanding benefits that include health, dental, and vision plans that begin on your first day and a retirement plan that is one of the best in the nation. You also receive 13 paid holidays and 15 vacation days beginning your first year. We also offer 2/3 tuition reduction at UCLA. Qualified candidates will have 3–5 years of prior healthcare accounting experience, including preparation of financial statements and journal entries, as well as excellent analytical and problem-solving skills. Skill in analyzing financial information and a thorough knowledge of accepted accounting principles and practices are also required. Bachelor's degree with a major in Accounting, Business Administration, or closely related field is required. For more information, please contact Reggie Glynn at: RGlynn@mednet.ucla.edu, Phone: (310) 794-0506; or apply online at: http://hr.healthcare.ucla.edu, and reference Job Code H52734. EOE.

635 UNIVERSITY OF WISCONSIN–LA CROSSE, College of Business Administration (AACSB-accredited) invites applications for two tenure-track accounting positions in its Department of Accountancy starting August 2011. Candidates seeking the rank of Assistant or Associate Professor should possess a Ph.D. or D.B.A. in Accounting (or a related discipline). ABDs will be considered at the instructor level until completion of their degree. Preference given to candidates with primary specialization in Tax or Assurance Services, with other areas seriously considered. For the tax position, a J.D. and advanced tax degree or extensive tax experience will be considered. Professional certification is highly desirable. These positions offer an array of resources for professional development. The college values continuing self-development in teaching, scholarship, and service consistent with AACSB standards. Review of applications will begin November 15, 2010 and continue until positions are filled. For additional information about this position and to apply, please visit: https://employment.uwlax.edu/. UW–La Crosse is an Equal Opportunity Employer committed to increasing the diversity of its workforce.

636 YORK UNIVERSITY, Schulich School of Business is seeking an outstanding scholar to fill a named Chair in Accounting at the rank of Professor. The appointment will be effective for July 1, 2011. Preferred candidates will have: a strong record of quality research in accounting, including financial accounting, managerial accounting, capital markets, international accounting, and/or capital markets research relating to accounting issues; a leadership position within academic accounting research marked by substantive and methodological innovation; strong evidence of teaching effectiveness and a willingness to employ a variety of sound pedagogies, including case-method teaching, that are consistent with the needs of the Accounting Area within the Schulich School of Business at both graduate and undergraduate levels, and eligibility for prompt appointment to the Faculty of Graduate Studies; willingness and ability to contribute to the doctoral program as a dissertation supervisor; a doctorate in accounting or related field. Interested candidates should submit: letter of application, including a statement of current research; curriculum vita; evidence of effective teaching; and three letters of reference. Please apply to: Dr. Thomas H. Beechy, Professor Emeritus of Accounting, Search Committee Chair, Schulich School of Business, York University, 4700 Keele Street, Toronto, Ontario, Canada, M3J 1P3; Email: tbeechy@schulich.yorku.ca; Phone: (416) 736 5097, (647) 588-3065. Deadline for applications: December 1, 2010. Salary and benefits are competitive. All York University positions are subject to budgetary approval. York University is an Affirmative Action Employer. The Affirmative Action Program can be found on the York University website at: http://www.yorku.ca/acadjobs, or a copy can be obtained by calling the Affirmative Action office at (416) 736-5713. All qualified candidates are encouraged to apply; however, Canadian citizens and permanent residents will be given priority.

637 BARUCH COLLEGE–CUNY, Stan Ross Department of Accountancy invites applications for anticipated openings for tenure-track and visiting positions at all ranks (Assistant Professor, Associate Professor, Full Professor, Clinical Professor, Distinguished Lecturer, Instructor, and Lecturer). Areas of specialization include financial and managerial accounting, taxation, auditing (internal and external), and accounting information systems. Rank depends on qualifications, and salaries are competitive. For ranks of Assistant/Associate/Full Professor, candidates should have a completed Ph.D. degree, a record of excellent scholarly research and demonstrated teaching effectiveness. Please submit curriculum vitae, a research paper, and letters of reference to: Professor Igor Vaysman, Stan Ross Department of Accountancy, One Bernard Baruch Way, B12-225, New York, NY 10010. Positions are open until filled. Baruch College is an AA/EO/IRCA/ADA Employer.

638 YORK UNIVERSITY, Schulich School of Business invites applications for a tenure-stream position in accounting at the Senior Assistant, Associate, or Full Professor level, effective July 1, 2011. Preferred candidates will have: a leadership position within the discipline, marked by significant contributions to accounting or auditing research; a track record of publications in top-tier journals; evidence of an ongoing research program with potential for further publications in top-tier journals; a willingness and demonstrated ability to supervise masters and/or doctoral students, and eligibility for prompt appointment to the Faculty of Graduate Studies; evidence of effective teaching and a willingness to employ a variety of sound pedagogies consistent with the needs of the Accounting Area and the Schulich School of Business; a doctorate in accounting or related field. The successful applicant will be expected to support both undergraduate and graduate education. Please apply via email to: Dr. Cameron Graham, Chair of the Accounting Search Committee, Email: cgraham@schulich.yorku.ca. Applications should include: an application letter, a curriculum vitae, three letters of reference, a current working paper, a one-page statement of current research program/interest and information about teaching performance. Deadline date for receiving applications is December 1, 2010. Salary and benefits are competitive. All York University positions are subject to budgetary approval. York University is an Affirmative Action Employer. The Affirmative Action Program can be found on York's website at: http://www.yorku.ca/acadjobs, or a copy can be obtained by calling the Affirmative Action office at (416) 736-5713. All qualified candidates are encouraged to apply; however, Canadian citizens and permanent residents will be given priority.

639 COLUMBIA UNIVERSITY, Business School seeks applications for faculty positions in Accounting. Rank is open. Appointment at senior rank requires outstanding record of scholarship. Candidates with an interest in Financial Accounting research will be preferred. We are looking for an outstanding person with evidence of a strong commitment to both teaching and research. Candidates should possess a doctoral degree or be close to completion. We offer a highly supportive research environment and easy interaction with faculty in other disciplines. Salary and benefits are competitive. Please send resume, an unpublished research paper, and letter(s) of recommendation to: Professor Tim Baldenius, Columbia University, Business School, 617 Uris Hall, 3022 Broadway, New York, NY 10027. To apply for this position, go to our online application site: http://academicjobs.columbia.edu/applicants/Central?quickFind=52217. Deadline: January 2, 2011. Columbia University is an Equal Opportunity/Affirmative Action employer.

640 COLLEGE OF SAINT BENEDICT/SAINT JOHN'S UNIVERSITY is seeking to fill two tenure-track positions in Accounting and Finance, at the rank of Assistant Professor, to begin August 2011. The successful candidates will be expected to teach upper- and lower-division accounting courses in our curriculum. Members of the department occasionally share in the teaching of courses in the common curriculum. All members have a commitment to supporting student research. The ability to teach one or more of the following courses would be an advantage: Information Systems, Audit, or Advanced Accounting. Preferred qualifications include Ph.D., D.B.A., J.D. with underlying accounting degree, or Master's Degree in Accounting/Business/Finance plus CPA certificate. If preferred qualifications are not met, the minimum requirement is a Master's Degree in Accounting, Business Administration, Finance, or Taxation. For further information and apply online visit: https://employment.csbsju.edu. Affirmative Action/Equal Opportunity Employer.

641 UNIVERSITY OF MASSACHUSETTS AMHERST, Department of Accounting seeks a collaborative leader as Department Chair who will work to further enhance the Department's stature as a leading accounting program, develop strategies for undergraduate and graduate education, and strengthen relationships with an extensive alumni network. Applicants should have a record of accomplishment worthy of tenure and appointment as Full Professor at a research-intensive university. Candidates with leadership experience and significant roles interacting with university and corporate stakeholders are preferred. Please submit a curriculum vitae, contact information for three references, and copies of research papers or publications to: Dr. Chris Agoglia, Chair, Search Committee, Room 222B, Isenberg School of Management, University of Massachusetts, 121 Presidents Drive, Amherst, MA 01003. Application review will start November 15, 2010 and continue until the position is filled. The University of Massachusetts Amherst is an Affirmative Action/Equal Opportunity Employer. Applications from women and minority candidates are especially encouraged.

642 ARKANSAS STATE UNIVERSITY, College of Business is recruiting applicants for the position of Assistant Professor of Accounting. The position is a tenure-track, 9-month position engaged in teaching, research, and service and requires a terminal degree from an AACSB-accredited university. (ABDs near completion will be considered.) Please visit: https://jobs.astate.edu for detailed information and to apply for position F00440. Applications will be accepted through October 1 or until the position is filled. Job Duties: Teach various accounting courses as assigned, engage in research leading to publications in peer-reviewed journals, perform service activities at the department, college, and university level. Technical areas most desired are: Auditing, AIS, Financial, and Managerial. Arkansas State University is an Equal Opportunity/Affirmative Action Employer with a strong institutional commitment to the achievement of excellence and diversity among its faculty and staff. ASU is also committed to creating a productive workplace in which both persons and property are secure. To achieve that goal, background investigations are conducted on all final applicants recommended for employment.

643 EASTERN CONNECTICUT STATE UNIVERSITY is seeking applicants for a full-time, tenure-track appointment in the field of Accounting at the Assistant/Associate level, to begin August 2011. Applicants should hold a Ph.D. in accounting or a related business discipline or be ABD with expected completion by May 2012. Candidates must demonstrate a commitment to teaching, scholarly, and professional activities, and a willingness to be active in service to the college and community. Applicants should hold a Ph.D. in accounting or a related business discipline or be ABD with expected completion by May 2012. Eastern is a co-educational public liberal arts university serving approximately 5,400 students. Additional information may be found at: http://www.easternct.edu. Applicants should submit a statement of interest, current vita, graduate transcripts, and three letters of reference to: Dr. Richard Silkoff at: silkoffr@easternct.edu.

644 UNIVERSITY OF MICHIGAN ANN ARBOR, Stephen M. Ross School of Business has tenure-track positions available in Accounting starting in the 2011 academic year. Candidates with research and teaching interests in all areas of accounting will be considered. Faculty positions at all ranks are available. The application should include a research paper, which may be based on the candidate's dissertation, evidence of teaching experience (if any), and a curriculum vitae that includes three references. The University of Michigan offers competitive salary and fringe benefits. A qualified junior candidate must have a Ph.D. or be in the final stages of their dissertation. A qualified candidate must also demonstrate the ability and interest to conduct high-quality, scholarly research as well as a high level of teaching competence. A qualified senior candidate must have a Ph.D. and a proven record of high-quality research, as evidenced by publication in the top accounting journals. A qualified candidate will also have a demonstrated record of teaching competence and evidence of the ability and willingness to contribute to school and professional service activities. To apply, please visit the Stephen M. Ross School of Business at the University of Michigan website: http://www.bus.umich.edu/FacultyRecruiting/. The review of applications will begin on November 30, 2010, and we strongly encourage all applications to be completed by that date. Applications will continue to be accepted until January 15, 2011. The University of Michigan is an Equal Opportunity/Affirmative Action Employer.

645 UNIVERSITY OF DETROIT MERCY, College of Business Administration seeks two tenure-track Assistant Professors of Accounting. Emphasis on teaching excellence; publications and service also required. Applicants should possess a doctorate in accounting or be ABD. Preference given to expertise in auditing/forensic accounting. More information is at: http://www.udmercy.edu/hr/jobs/. Send a letter of interest, resume, contact information for four references, and evidence of teaching effectiveness (if available) to: Dr. Jeanne David at: jeanne.david@udmercy.edu. Review begins immediately and continues until positions are filled. Michigan's largest, most comprehensive private university, the University of Detroit Mercy is an independent Catholic institution of higher education sponsored by the Religious Sisters of Mercy and the Society of Jesus. The University seeks qualified candidates who will contribute to the University's mission, diversity, and academic excellence. The University of Detroit Mercy is an Equal Opportunity/Affirmative Action Employer with a diverse faculty and student body, and welcomes persons of all backgrounds.

For focused results in accounting and tax research, start here.

While the world of accounting information continuously transitions, professionals, educators, and a growing number of accounting students look to reliable resources for researching accounting trends, rules, issues, and history pertaining to the future of accounting.

ProQuest Accounting & Tax, our premier accounting research database, makes research time more effective with centralized access to top accounting journals, reference reports, proceedings, dissertations, and more.

Researchers can instantly cross-search over 2,300 publications, with hundreds in ongoing full text, plus dissertations and working papers covering accounting and tax topics.

ABI/INFORM subscribers can enhance their business database library with ProQuest Accounting & Tax, which provides a vast array of additional full-text accounting journals, trade publications, dissertations, and working papers not included in ABI/INFORM.

ProQuest Accounting & Tax includes:

- Highly ranked global and scholarly journals, including all journals from the American Accounting Association
- Accounting, tax management, and auditing trade publications from top publishers such as CCH, Aspen, Thomson Professional and Regulatory Services, Source Media, Tax Management Inc., and Euromoney
- AICPA publications
- International Federation of Accountants (IFAC) standards
- Tax law reviews from top institutions such as New York University, Stanford, Vanderbilt, and University of Virginia
- Tax law and financial law publications from noted sources such as the American Bar Association and American Law Institute
- Global accounting and tax journals from the UK, Canada, Australia, France, Spain, Ireland, and more
- Financial advising, governance, and risk management publications from Incisive Media, CFA Institute, and Association for Financial Counseling and Planning
- Accounting history from publications such as *The Accounting Historian's Journal*
- Over 1,500 accounting dissertations and more than 7,800 working papers from SSRN, plus select conference proceedings



ProQuest Accounting & Tax

Visit **www.proquest.com/go/accounting** today to learn more, and sign up for a free trial!

ProQuest® Start here.

UNITED STATES POSTAL SERVICE®

Statement of Ownership, Management, and Circulation (All Periodicals Publications Except Requester Publications)

1. Publication Title	2. Publication Number	3. Filing Date
The Accounting Review	0 0 4 _ 0 0 0	9/24/10
4. Issue Frequency	**5. Number of Issues Published Annually**	**6. Annual Subscription Price**
6 times a year in January, March, May, July, September, and November	6	$375

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4®)*

5717 Bessie Drive, Sarasota, Florida 34233-2399

Contact Person: Lisa Habblitz

Telephone *(Include area code)*: 941-556-4115

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

5717 Bessie Drive, Sarasota, Florida 34233-2399

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*

American Accounting Association
5717 Bessie Drive, Sarasota, Florida 34233-2399

Editor *(Name and complete mailing address)*

Steven J. Kachelmeier, The University of Texas at Austin, McCombs School of Business, Dept. of Accounting, CBA 4M.202, 1 University Station B6400, Austin, TX 78712

Managing Editor *(Name and complete mailing address)*

Same as above

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
American Accounting Association	5717 Bessie Drive, Sarasota, Fl. 34233-2399

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

Full Name	Complete Mailing Address

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*

The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:

☒ Has Not Changed During Preceding 12 Months

☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

28801 1

13. Publication Title	14. Issue Date for Circulation Data Below
The Accounting Review	July 2010

15. Extent and Nature of Circulation			Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies *(Net press run)*			3,831	3,689
b. Paid Circulation *(By Mail and Outside the Mail)*	(1)	Mailed Outside-County Paid Subscriptions Stated on PS Form 3541(Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)	3,421	3,441
	(2)	Mailed In-County Paid Subscriptions Stated on PS Form 3541 *(Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)*	—	—
	(3)	Paid Distribution Outside the Mails Including Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Paid Distribution Outside USPS®	—	—
	(4)	Paid Distribution by Other Classes of Mail Through the USPS (e.g. First-Class Mail®)	—	—
c. Total Paid Distribution *(Sum of 15b (1), (2), (3), and (4))*			3,421	3,441
d. Free or Nominal Rate Distribution *(By Mail and Outside the Mail)*	(1)	Free or Nominal Rate Outside-County Copies included on PS Form 3541	6	6
	(2)	Free or Nominal Rate In-County Copies Included on PS Form 3541	—	—
	(3)	Free or Nominal Rate Copies Mailed at Other Classes Through the USPS (e.g. First-Class Mail)	—	—
	(4)	Free or Nominal Rate Distribution Outside the Mail *(Carriers or other means)*	—	—
e. Total Free or Nominal Rate Distribution *(Sum of 15d (1), (2), (3) and (4))*			6	6
f. Total Distribution *(Sum of 15c and 15e)* ▶			3,427	3,447
g. Copies not Distributed *(See Instructions to Publishers #4 (page #3))* ▶			404	242
h. Total *(Sum of 15f and g)* ▶			3,831	3,689
i. Percent Paid *(15c divided by 15f times 100)* ▶			99%	99%

16. Publication of Statement of Ownership

☒ If the publication is a general publication, publication of this statement is required. Will be printed in the November 2010 issue of this publication. ☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner	Date
Tracey Sutherland Executive Director	9/24/10

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

THE ACCOUNTING REVIEW
ISSN 0001-4826

PERIODICALS POSTAGE PAID
SARASOTA FLORIDA
AND ADDITIONAL MAILING OFFICES

POSTMASTER
Send Address Changes To:

AMERICAN ACCOUNTING ASSOCIATION
5717 BESSIE DRIVE
SARASOTA, FL 34233



**********3-DIGIT 482 ***
P-2 P106
14918 REVIEW 1
MARYGROVE COLLEGE
LIBRARY SERVICES
8425 W MCNICHOLS RD
DETROIT MI 48221-2546

1023